

**Parapsychology, a critical and experimental
study with special reference to psychokinesis
and to problems of methodology and interpret-
ation of results.**

by

H. H. J. Keil, B.A. (Hons.)

**Submitted in fulfilment of the requirements
for the Degree of Doctor of Philosophy.**

University of Tasmania.

To the best of my knowledge the material presented in
this thesis is original and has not been submitted for examination
purposes previously unless due references are given.

Hobart, Tasmania, March, 1964.

(Sgd.)

H. H. J. Keef

Preface

In parapsychology more than in other fields of inquiry, it has been customary by some of the most significant contributors to present their writing in the first person. When I restrict this custom to this preface, it is certainly not in disrespect to authors like Tyrrell and Rhine but in recognition of the formal aspects of this thesis.

However, avoiding the first person does not mean detachment from the issues involved and my admiration for clear detached writing ranks high for Bridgman's contributions (1959) to less controversial fields in which he wrote in the first person, probably with the deliberate intention of demonstrating to the reader that the human recorder of experiments is part of the experimental set-up which he describes.

It may therefore be appropriate to list here in the way I see it, certain aspects of my personal bias.

When I reached some understanding about probability statements, I tried, and I believe I succeeded, in incorporating probability into my general outlook on life.

I tried to replace certainties by probabilities which change in magnitude as relevant information becomes available, without losing sight of the necessity to make decisions on the basis of probabilities.

My first introduction to the field of parapsychology came about through a highly successful and open minded professor of physics and head of a university department and I have found in this man and in other scientists who were not directly connected with parapsychology more inspiration than in some parapsychologists of high standing.

Such experiences led to a high probability estimation as to the value of attracting interested although to some extent sceptical scientists to the field, at least as long as there is no "breakthrough" in the control of parapsychological events.

Most of these scientists were orientated along the line of physical monism. But they did not "swallow babies for breakfast" (p.¹19).

Five years ago my probability estimations of dualism and physical monism were about 0.5 each but

1. Page references which are underlined refer to this thesis.

since then I began to appreciate to some extent the enormous amount of change that has taken place in basic scientific concepts during the last 60 years or so, and this change is still continuing.

It seems to me that the kind of physical monism that was recognised 60 years ago is now so far removed from our present day knowledge of the physical world that the kind of different events on which dualism and monism were based 60 years ago have the same order of difference as the kind of events on the basis of which monism developed during the last 60 years.

The sort of thing I am trying to point out here was also discussed by Stevens and Rhine (Stevens, 1930) and I find myself much more in agreement with Stevens than with Rhine.

It seems to me rather doubtful whether it is useful to speak of physicalism today, because it is more often than not mistaken for or associated with, the physicalism of 60 years ago and it seems equally or more doubtful whether dualism can be seen as a real alternative to present day physicalism.

A kind of duality has been supported in physics in connection with the principle of complementarity

but there are some doubts as to how far this support is justified (Landé, 1959).

My own experiences suggest that any advantages which can be gained by emphasising what appears to me only a possibility - dualism on the basis of parapsychology - are in the long run small, compared to the disadvantage of losing scientists who might otherwise be interested in participating in research and discussion.

At the same time I believe I can see that research in the field of parapsychology may at times have only survived because of support from donors with a dualistic outlook based on religious or other beliefs.

In my personal estimation, progress in parapsychological research has been painfully slow. If my view is acceptable and if it is not considered unduly pessimistic to expect more years of hard labour with little prospect of a real breakthrough ahead, then the participation of more qualified scientists is an urgent matter indeed. Compared to the highly complex basis of present day physics I do not consider that withdrawing statements about the non-physicality of parapsychological events would be a kind of cowardly act in order to be more acceptable to the fashion shows of our time.

(v)

There is no need to accept any form of physicalism as more than another possibility. But to advertise parapsychological events as non-physical with a great amount of conviction seems to me an emotional reaction to the scientific fashions of the beginning of this century.

Parapsychologists should be specialists in their field but what special technical knowledge is necessary can be acquired by any competent scientist with psychological training and yet what makes a successful parapsychologist is still a rather mysterious and untrainable quality. As long as we do not know anything more precise about this it seems to me highly undesirable to form a kind of exclusive club and to reject anybody who did not reach the 0.01 level of significance. I have some doubts whether the total evidence justifies any such divisions, but having entered this field myself without any previous spontaneous experiences, I find it hard to believe that the kind of parapsychological events (unless all my results were due to Type 1 or other errors) which I recorded, should be restricted to the peculiarities of my personality.

Some other personal considerations which have a bearing on this thesis may also be included here.

When I commenced work for this thesis by planning and carrying out experiments in the field of parapsychology I had no personal contact in Australia with other workers in this field. I had met Dr. Pratt, who was until recently Deputy Director of the Parapsychology Laboratory, Duke University, some years earlier in Germany, but apart from this single discussion during one afternoon I have had no opportunity to speak with other psychologists.

A.L.

Professor McAuley, who was head of the Physics Department of the Tasmanian University and to whom I have referred earlier, had retired for health reasons and had withdrawn from all University activities.

After working for 9 to 12 months I began to feel the stress of isolation and began to look for any possibility to work with parapsychologists overseas. It seems to me now that this was a necessary step which I had to take although my departure to the USA interrupted in a sense some of the experimental research which I had carried out in Tasmania.

I had planned to repeat certain aspects of my

work in the Parapsychology Laboratory, Duke University, USA but I found more technical difficulties there than I had anticipated. I would like to mention here that Dr. Pratt had questioned prior to my arrival there whether it was advisable to repeat certain experiments in a new situation and he had pointed to the difficulties in setting up apparatus designed for a different voltage.

It was therefore on my own responsibility that I tried to continue with some aspects of the repetitions but after having worked in a comparatively small psychology department in one of the smallest universities I did not expect the technical difficulties in obtaining tools and apparatus with which I became confronted. As a result of these and also because of long delivery times for some necessary small items it was for some months uncertain whether I would be able to reconstruct the PK apparatus which I had used in Tasmania within a reasonable time.

I was eventually able to reconstruct the main section of the apparatus and to demonstrate to the members of the Duke Laboratory, however the experimental sessions were of doubtful value because without a really satisfactory base for mounting the apparatus, the

automatic distribution tended to drift away from the 50/50 distribution which was set up prior to the experimental sessions and which had remained more or less unchanged during the earlier Tasmanian experiments.

The sort of difficulties described here are well known to anyone actively engaged in research involving apparatus. Nevertheless it is hoped that these more personal remarks will be of some help in providing the reader with some background information to the experimental reports presented here.

I would like to use this opportunity to express my gratitude for his most stimulating influence on my earlier work, to Professor McAulay, who as head of the Physics Department of the University of Tasmania, introduced me to the problem of Parapsychology.

I appreciated very much the patience, tolerance and encouragement which I received from Professor Cardno, Head of the Psychology Department, University of Tasmania, who supported my application to carry out research in parapsychology for this PhD. thesis. My work was further assisted by the laboratory facilities and equipment of his department.

I would like to thank Dr. Rhine and Dr. Pratt who recommended a research fellowship from the Parapsychology

(ix)

Laboratory, Duke University, which together with a post graduate research fellowship from the Duke University enabled me and my family to stay for 8½ months at Duke. Dr. Rhine and Dr. Pratt provided a most valuable link with the - from the Tasmanian point of view - outside world of parapsychology.

Mr. Hasofer, lecturer in Mathematics, University of Tasmania, was most helpful in suggesting the tests for randomness and in checking the necessary calculations.

I was fortunate to be able to discuss parapsychological questions with Dr. Tenny, Dr. Freeman, Mr. Saleh, Mr. Roll, Mr. Cox and Dr. Kewley at Duke, with Dr. Osis and Dr. Dean in New York and with Professor Bender and Professor Tenhaeff in Europe. I was also fortunate to meet Professor Thouless when he visited Australia.

Dr. Tenny's arrangements which enabled me to participate in EEG work and parapsychological tests at the Veterans' Hospital, Durham, N.C. are also gratefully acknowledged.

I would like to thank Mrs. Langworthy for her careful and patient check through a large proportion of the original data and summaries and for typing this thesis.

Last but not least my sincere thanks go to all those students who participated in research.

Content

Preface	1
Introduction	1
A short historical note	8
Part I : Problems in Psychology	
The meaning of parapsychology and methodological problems	27
Emotions and 'surplus' meaning in science and parapsychology	63
The problem of non-physicality	78
Numbers as targets	102
Summary	108
Part II : Experimental report	
Introduction	111
A GESF group experiment disguised as a subliminal perception test	123
A GESF experiment disguised as a subliminal perception test with individual test sessions and electric shocks as reinforcements	150
The PK apparatus	172
Additional details about PK apparatus	192
Tests for randomisation	201

Recording errors	204
The PK apparatus experiment	208
Some considerations about future work with the PK apparatus	236
The Duke experiment with the PK apparatus	244
A standard PK test with dice	248
Physiological measurements and parapsychological processes	253
A GESP music test with favourite music targets	264
Short summary of experimental reports and future work	284
References	287
Appendices	305

Parapsychology

A critical and experimental study with special reference to psychokinesis and to problems of methodology and interpretation of results.

Introduction

Parapsychology may be defined as a "Grenzgebiet" or fringe area of psychology dealing with events that are associated with, or are part of, behaviour and which do not fall within the scope of contemporary general psychology.

This definition is by no means a satisfactory one and some attempts will be made later to clarify parapsychological terms (pp. 91-93). However, as it stands, this definition does not introduce the problem of non-physicality.² Rhine's arguments

2. Although Thouless and Wiesner (1948) did not introduce 'psi' as physical or non-physical, this term has been initially avoided here because it is associated in the Journal of Parapsychology Glossary (1963) with non-physical. (This was not the case in 1950). Other terms which are used here without further definition correspond to the terms as defined in the Glossary. However, the older definitions in the Glossaries (e.g. 1950) are sometimes preferred to more recent ones (1963) but the differences are not sufficiently large to warrant a detailed discussion here. But note, for instance, the difference in the definitions of ESP and PK as printed in 1950 and 1963.

for the non-physicality of parapsychological processes (1957, p.247) will also be discussed in some detail.

The majority of experiments carried out in this field during the last 20 years or so did not follow any long term master plan but were of an exploratory or short term nature. This seems to be characteristic of what might be called the pioneering stage. Both Murphy (1938) and Rhine (1962) came to the conclusion that long term planning may now be a desirable step towards further progress. The increased number of research institutions and the more generous financial support which has been forthcoming recently (Parapsychol. Bull., 1962, No. 61; 1963, Nos. 64 & 66), add reality to such proposals.

This thesis represents a very modest attempt that has been made to carry out experiments along the lines of an overall plan.

While it cannot be possibly claimed that a certain approach to parapsychology has been fully explored experimentally, it is hoped that the work presented here may help to draw some tentative conclusions and to identify within a limited area, problems which have not been dealt with so far.

The approach has been to combine a number of factors which are believed to be favourable towards parapsychological results and to work, generally speaking, towards the goal of less unstable test conditions.

To this end experimental conditions were considered and tested in which the subjects were not aware of the problematic aspects of the experiments, and the reinforcement of correct responses was attempted by introducing electric shocks as "negative" stimuli and by the projection of slides as "positive" stimuli. "Correct" and possibly parapsychological responses resulted automatically in the presentation of positive stimuli and incorrect responses in negative stimuli. In some experiments the complete set-up including the scoring of responses, was automatic.

One apparatus produced automatically, a distribution which was found to be random within the limits of randomness tests. A basis of comparison was thus provided, not so much in response to some probably unjustified criticism (Girden, 1962), but in order to include the possibility of detecting less definite parapsychological processes, e.g. if such processes are present but operate in the wrong direction (e.g. Carington, 1940a, 1940b ; Reeves & Rhine, 1943 ; Bindrim, 1947 ; Cadoret & Pratt, 1950)

3. References will be headed 'e.g.' if those quoted represent a large number of possible references. The references actually selected are those which seem representative and which were available to the author. If the original publications were not available or could not be obtained within a reasonable margin of cost and inter-library loan service, references to later reviews or discussions are made.

then it may still be possible to detect differences between parapsychological processes and random processes if random processes can be provided for comparison.

In undisguised tests subjects also used their favourite pieces of music as targets in a GESP experiment. This was an attempt to stabilize the test situation by providing the subjects with individual targets of personal interest.

Techniques in measuring physiological variables during parapsychological tests were also investigated.

The approach described here should be regarded as a possible one rather than as the desirable one. Little consideration has been given to personality differences which could provide another major line of approach which has been explored for some time elsewhere (e.g. Stuart et al., 1947 ; Schneidler, 1949 ; Humphrey, 1950), but it is argued that one ought to try and see whether, under sufficiently strong stimulus conditions, a parapsychological response can be elicited irrespective of the personality factors of the subjects and of the experimenter.

It seems however that the psychological variables introduced were not as strong as one might like to see them ^{test conditions} under because volunteer students cannot be expected to endure severe electric shocks, yet a mild shock may not serve as a negative stimulus, and

it is difficult to agree on any positive stimuli. Nevertheless the results suggest that even this sort of approach can be successfully employed and that further work should be carried out.

In choosing this particular line of approach in the experimental investigations the author was aware that obtaining significant test results in a particular setting does not necessarily provide much help for the field of enquiry.

Murphy's critical evaluation of contemporary experiments (1958) has relevance to the experiments presented here.

However to carry out research not only in an unorthodox field but also in a somewhat unorthodox manner within the field seemed too ambitious, allowing for the limitations of time and material.

Kahn's analysis (1962) of historical events which lead to breakthroughs in scientific research is a reminder that the problems one attempts to solve are often not the problems which are relevant to the field, but also that an extraordinary mind is needed to reformulate the crucial problems.

Reports about parapsychological work in Russia (Ryal, 1961a, 1961b ; Pratt, 1962) are difficult to analyse but it appears that the sort of interest and discussion that goes on there among many prominent scientists is the kind of thing that is needed in the West. And if this is correct then there seems to be at present one real problem in parapsychology that can be

identified with some confidence, that is, the lack of interest and participation in research and discussion by scientists in the West.

It seems to this author that the Western parapsychologists are not without blame for this state of affairs and some theoretical considerations question the assumptions and interpretations of evidence which may have brought about this situation.

On the experimental side it is hoped that the sort of approach developed here may lead one day to a kind of 'test kit' which could be handed out to interested scientists and which contains test material but also tape recorded instructions, perhaps some special motion pictures and sound records to create the correct emotional atmosphere, perhaps ^a personality test on the basis of which parapsychological results can be extracted in more concentrated form, and automatic scoring devices. All this should enable the scientist to remain sufficiently dissociated with the experiment to exclude him as a variable in the test situation.

It is not certain whether even a partial solution to this problem can be achieved by the experimenter's dissociation from the subject's test situation. West and Fisk (West, 1936) obtained results which suggest that a rather remote connection

(unknown to the subjects) can still influence the outcome of parapsychological events. Yet without denying the difficulties and complexities of this problem it seems reasonable at least to attempt to remove the experimenter from the subject's parapsychological 'field'.

The tests may also be disguised to minimize the conscious or unconscious effects which subjects may have on the parapsychological events because of the controversial nature of the experiments.

How far the work here can be regarded as a positive contribution towards the solution of this problem shall be left unanswered but it seems to this author that the investigations which were attempted are the kind of steps which among others ought to be taken at this stage.

In discussing basic problems in methodology in connection with this parapsychological work, it is hoped that some of the problematic aspects of parapsychology will be somewhat clarified, but it is also hoped that the clarification of problems in a fringe area of psychology has some bearing on the methodological basis of general contemporary psychology.

A short historical note

This section will have to remain extremely limited not only⁴ because of the difficulty of obtaining relevant publications but also because of the near impossibility of deciding what historical material can be regarded as parapsychological.

To avoid too personal an interpretation of such material the main attempt will be to see whether historical considerations can be used to fit the present day experimental findings into a wider frame-work.

Discussions on orthodoxy (I. B. Cohen, 1952 ; Boring, 1952) show that the calculated probability that an event took place must be judged against the background of other available knowledge about the kind of alleged events. History is part of this background.

In the Parapsychological Bulletin (No. 60, 1962) it was reported that Vasiliev, a well known Russian physiologist who had worked under the late ^{V.H.} Bekhterev, conducted experiments in which he apparently influenced the hallucinatory experiences of peyote drugged subjects with a big horse shoe magnet.

4. The relevant literature available in Tasmania is very restricted indeed. Specific publications could be obtained through inter-library services but these were insufficient for a historical survey.

Close to 200 years ago Mesmer reported similar results with hypnotised subjects (Goldsmith, 1934 ; Zweig, 1936). Perhaps this completion of a historical loop appears at first sight amusing if it were not for the somewhat sobering thought that it took a very long time to investigate by comparatively simple means, alleged events, which at least since the discovery of electric activities in the living brain could remain in the framework of orthodox scientific thought.

It is possible to argue at great length against the dependence of scientific research on the changing fashions and beliefs. It shall here only be summed up by one example given by I. B. Cohen (1952, p.507).

"We may note a curious historical phenomenon with regard to "scientific respectability" in our attitude towards ancient and mediaeval atomists and alchemists. Practically every book on science, or even the history of science, hails the atomists (such as Democritus or Lucretius) as precursors of our present science, even though modern Physics demonstrates that their notions were primitive and unlike our own. But the alchemists, whose notions have been shown by modern chemistry and physics to be in some degree "correct," are not characterized as respectable precursors. (For instance, Charles A. Browne wrote an essay on "Error in chemistry" (The story of human error, ed. by Joseph Jastrow, N.Y., Appleton-Century Company, 1936) in which he praised the

ancient atomists extravagantly ("Had the atomic theory of Democritus met with immediate acceptance, the rise of modern chemistry might have been advanced two thousand years") and castigated the alchemists (" . . . a pseudo-science . . . , " . . . The fatuity of alchemy . . . , " etc.) Here we probably exhibit an inherited prejudice of the nineteenth century and our views of ancient atomism and alchemy are reflections of the state of nineteenth-century science which held that atoms are the fundamental indivisible parts of matter and that chemical elements are fixed and can under no circumstances be "transmuted" into other chemical elements."

One conclusion may be drawn which is :

New discoveries, particularly if they are not clearly understood, are of little value to science unless they are presented in a way in which the disagreement with the accepted background of relevant research is not any wider than the new findings necessitate.

In modern parapsychology the postulate of non-physicality may be one of the unnecessary stumbling blocks which stand in the way of wider acceptance and participation in research.

Perhaps even here Mesmer may be taken as a somewhat tragic example of a man who dedicated his life to the discovery of a new 'Grenzgebiet' but who failed, as far as can be judged now,⁵

5. Tenhaeff (1937, pp. 73-74) in a careful analysis agreed that no evidence for any kind of 'fluidum' can be presented at present but he does not exclude this possibility entirely.

to penetrate to the more essential features of his discovery by emphasizing a certain interpretation which was not a necessary one.

This should not be understood as a submission of facts to fashions but as an attempt to check one's own interpretations more carefully when they are made in a new border area.

However the history of parapsychology did not start with Meemer but can be traced back to time immemorial. It is a matter of personal choice to conjecture just what sort of magical practices may have had a parapsychological element. The Oracle at Delphi may have predicted the future mainly on the basis of facts known at the time of enquiry but at least the possibility that some of the more remarkable announcements may have been due to parapsychological factors cannot be ruled out. At any rate a kind of hypnotic state was more than 2000 years ago apparently considered helpful in the attempt to predict the future (Hyslop, 1906).

Various phenomena in connection with some religious miracles may be interpreted in parapsychological terms. This is^a difficult task because no reliable evidence exists for similar phenomena under controlled conditions unless one is prepared to accept for instance the accounts of levitation of the medium Home (Myers, 1903, p. 581) or Smythies impressive report (1951) of a phenomena which may or may not be called levitation.

Leaving out religious interpretations, similar events reported (Leroy, 1928) about saints may then not necessarily be seen as false or mistaken description by the witness present.

Levitation is no doubt an extreme example which is not likely to be accepted by many parapsychologists. Nevertheless some conclusions can be drawn.

Levitation (if it exists) as a parapsychological event is in agreement with similar events which are reported as occurring in certain religious settings. From the scientific control point of view both sets of events cannot be accepted as established. However if one set was accepted the background probability would become more favourable to the other. From a speculative point of view one may look for independent observations of both sets of events and consider how likely it is that remarkable similarities should have occurred. If it becomes likely then that such events were perceived by those who reported these events the question of misperception, illusion and hallucination must be raised and the various possibilities must be assessed on the basis of scientific knowledge available.

If it should be found after a detailed analysis - which will largely remain speculative but which may include experimental evidence on e.g. hallucinatory experiences - that there remain some elements in the reported events which are unexplained by orthodox theories and which are challenging, then these elements

are not acceptable as scientific facts but they may be of sufficient importance to change the background probabilities against which new experimental findings are judged.

The problem of non-physicality will be discussed from the experimental point of view (p. 78) but it may be of some interest to consider the historical events which lead to this problem and in particular to see where psychology stands now.

The historical development of the mind-body question need not be repeated here (Murphy, 1946) but it may be worth while to try to see why the kind of dualism developed in the Western world of thought and why this dualism became so important.

In the last analysis it is impossible to say whether differences between the Western and Eastern civilizations are based on accidental or random developments or on more systematic influences from the environment or from special characteristics of the peoples concerned.

The early advances in technology and science in the West may be regarded as accidental but the preference for analysis is perhaps a characteristic weakness and strength of Western thought.

Analysis may have brought about the particular major arguments for dualism but some of the meaning is added through the Western assumption of one body, one soul and/or one mind. It is only in recent years that the ideas of less clearly definable limits which merge into each other have become more respectable if still not acceptable.

In psychology Jung (1959) has probably contributed most to blurr the boundaries between single individuals but his writing is reflected to a small extent in the world of physicists like Heisenberg (1955), Jordan (1955), and Margenau (1957).

Nevertheless the majority of modern psychologists avoid any reference to the mind-body question. If this is seen as a reflection of the present 'Zeitgeist' the general interest in perceptual problems which have become a central theme in psychology may also be seen as an unconscious compensation for, or sublimation of, the awkward body-mind question.

After all perception seems to include elements which are irreducibly psychological. Sherrington (1951, pp. 248, 249) wrote:

"A star which we perceive. The energy-scheme deals with it, describes the passing of radiation thence into the eye, the little light-image of it formed at the bottom of the eye, the ensuing photo-chemical action in the retina, the trains of action-potentials travelling along the nerve to the brain, the further electrical disturbance in the brain, the action-potentials streaming thence to the muscles of eye-balls and the pupil, the contraction of them sharpening the light-image and placing the seeing part of the retina under it.

The 'seeing'? That is where the energy-scheme forsakes us. It tells us nothing of any 'seeing'. Much, but not that. A tiny patch of a particular radiant energy disturbing the surface of the body in a region specially reactive to it; it connects that patch with an energy-path entering the eye; then with one carrying brainward from it, a shower of repetitive electric potentials. It locates these in a certain region of the brain, which it therefore indicates as concerned with what occurs in us through the eye. It also accounts to us for all the manoeuvring of the eye-balls as they catch the photo-image and sharpen it and place the eye centrally under it, so too for our turning of the head to help the eyes.

But as to our seeing the star it says nothing."

The difficulty of thinking through the process of perception may at least be partly based on traditional semantic processes. Even if psychologists do not often mention mind as a separate entity in their text books it may be fair to assume that many still think as if they had individual minds somehow separate from their bodies. Yet the private perceptual worlds of individuals have sufficient similarity with each other to exclude the necessity of assuming separate minds.

There is little doubt that a good deal of perception is of a mechanistic nature and it is only when the sequence of impulses

which originate from the sense organs and which can be traced to specific centres in the brain start to convey a picture which we see, that conscious experiences begin which one traditionally reserves for higher animals and not for machines.

But if for a psychologist the study of behaviour (and the processes leading to it) is the primary task, then the 'perception' by a machine may lead to 'machine behaviour' which is difficult to distinguish from behaviour of an organism.

Although human behaviour is usually more complex and above all more flexible, comparatively complex operations are carried out automatically and are supervised by machines. The following example may illustrate this:

In a machine set-up a fuse blows out and the resulting breakdown of one circuit is transmitted to the control machine which may then be able to switch to an alternative circuit or restart the (automatic) fuse again after removing some overload.

Does the control machine see anything? In terms of behaviour the control unit received messages which were analysed correctly and which resulted in appropriate behaviour. If probability devices are built into the machine then it may be impossible to predict whether for instance the original blow out of the fuse will be followed by restarting the fuse or by the use of an alternate circuit.

So it will be difficult to reject 'machine perception' on the ground of human/^{un}predictability.

Perhaps all this machine behaviour can still be accounted for in a human being by reflex mechanisms which do not require any conscious responses. But if it is agreed that not all human behaviour is based on reflexes and if it is agreed that a particular kind of machine behaviour can still be accounted for in terms of reflex mechanisms at the human level then it may be possible to design more complex machine behaviour which goes beyond the reflex category.

Is the hesitation in speaking of the control unit (of the machine) as being conscious of a blow-out justified, or is it based on traditional thinking?

It is difficult to find a full logical justification for this hesitation if it is agreed that consciousness is not necessarily restricted to human beings (Sherrington, 1951, p. 220). In any order of lower to higher animals it is difficult to see where consciousness starts, yet if sufficiently simple organisms are considered to have some kind of consciousness then machine consciousness is not such an unreasonable term.

So perhaps the question is merely one of complexity.

Any utterances that the human being 'knows what he knows and what he does not know' does not provide a clear distinction either. Do the dolphins or higher apes 'know'? If there is some doubt, then there is little hesitancy in answering this question in the negative for somewhat lower animals. Even if it is only agreed

that the last events in the process of seeing a stimulus are machine-like up to say, a cat, then the arguments for special differences at the human level are weakened.

Obviously there is some gap between human beings and other higher animals. Perhaps the width of this gap is over emphasised in the West through religious traditions but even if it is as wide as some may argue (Buytendijk, 1958) some of the behaviour of retarded human beings may be compared with machine behaviour.

Is it certain that none of the perception of such human examples can be compared with machine perception?

If such comparisons are possible then one might proceed at the human level in small steps towards the normal human being and ask again at what point does a principal difference for the last stages of perception come in?

The above discussion is not aimed at making more acceptable the view that human perception is of the same kind as machine perception. But it may be argued that it is very difficult to single out any clear border between human and machine perception.

It may be precisely because machine behaviour has approximated human behaviour to an uncomfortable degree, that those who desire a clear distinction for religious or traditional or other reasons go to such length to reject monism. Koyle's opening remarks in *Man and Materialism* (1957) are amusing but also reflect those aspects of the 'Zeitgeist' which reject monism on other than logical grounds.

"What is a materialist? In a popular view I suppose a materialist is a pretty unpleasant person who gobbles babies for breakfast. This is a view I do not agree with. I am a materialist and I haven't gobbled any babies, yet. Nor has materialism anything to do with Soviet Communism. It is true that Communists prefer a crude style of materialism, but this has small similarity with the deeper materialism of the Western World."

It seems then, that as monistic interpretations explain an increasing area of events, the emotional reactions to these explanations increase also, and find expression in an urgent desire to demonstrate 'something else' which transcends monism and which is usually regarded to be on a higher and superior level.

It seems however, that the emotional reactions are largely unnecessary and based on a misunderstanding, or on an outdated view of material monism.

The material world as understood today has changed an enormous amount from the accepted view at the end of the last century.

Conant wrote 1951 (1961 ed. p.27) "There is no doubt about it, somewhere about 1900 science took a totally unexpected turn."

The complexity of material structure is so much more intricate and one might almost say mysterious (Oppenheimer, 1953) that there is now perhaps no need to feel disturbed about a material human being in the way one might have been disturbed about a crude automaton.

Indeed Broad suggested (1958, p.32) that perhaps a material world is "less disturbing place to live in". . . The world as it really is may easily be a far nastier place than it would be if scientific materialism were the whole truth and nothing but the truth."

From the very limited amount of secondary evidence about parapsychological events, it is difficult to decide whether these events fit into a material world or not; but before the question is answered in the negative, one has to consider the changes in the understanding of the physical world which occurred during the last sixty years or so and ask whether any isolated events which are accepted as physical now would have been accepted as physical sixty years ago if they had been known in isolation only. If one is not certain about the answer then the case for the non-physicality of parapsychological events is weakened.

There is also no sign that the present picture of material structure is complete. Indeed parapsychological events may suggest new fields and dimensions which however do not necessarily need to be fundamentally different from the rest of the material world.

Ryzi (1963) reported that Vasiliev suggested in 1962 that a new kind of energy may be detected wherever the highest order of material structure is present - that is, in the nervous system.

The tendency of apparently isolated sections of the brain to

start a rhythmic change of potentials which can be measured by EEG may indirectly point towards such new fields of energy.

Vasiliev's suggestion seems challenging enough without any expression of belief as to the non-physicality or otherwise of parapsychological events.

However on a more general level it is probably fair to argue that the less the observed events fit into an accepted framework of knowledge, the more general knowledge is likely to be gained from understanding these events. Electrical and radiation phenomena can be taken as an example where isolated kinds of events were known long before the whole phenomena were understood, at least to the extent that they fit now into an acceptable framework and can be manipulated to produce spectacular results.

Particularly in the case of electrical phenomena this process took a considerable time and it is perhaps a somewhat sobering thought to imagine some painstaking research by say, a Greek 2,500 years ago, trying to understand why a piece of amber rubbed with a cloth will attract and pick up small particles. Perhaps parapsychology is not in a much better position now and it is only because of the acceleration in scientific research during the last 300 years or so that one can be more hopeful.

To return to the difficulty of judging how meaningful historical reports are as a background to contemporary parapsychology

it may be of interest to consider for comparison one other area of investigation. Reports about unidentified flying objects (UFOs) have led to official investigations by various governments, to new beliefs, specific delusions, and to a vast number of speculations.

Historical records have also been reinterpreted to support specific speculations. Jessup's (1955) attempt in this direction was not a satisfactory one because references were not always given and some referred to periodicals which are not generally known for their scientific status. Nevertheless without checking the precise details, which does not seem necessary for the purpose of this discussion, one can see that Jessup was able to find a certain amount of support for his claims by searching through historical records.

Jessup used historical material to support two kinds of claims on which his belief in visits and contacts from space is based:

1. Reports of unidentified flying objects which are presented today are of a kind similar to those recorded at various times in history.
2. Certain ancient stone constructions - best known among these the Egyptian pyramids - cannot be satisfactorily explained in terms of present day knowledge about moving heavy stone-work. Consequently some intelligence from space must have played a part in it.

The support for the first claim is of limited value even if the historical records are clearly acceptable. If human observers see lights in the sky now and if these lights are not connected with intelligence from outer space, then it must be expected that similar lights were seen previously. The only improvement is introduced by excluding misperceptions from reflecting planes or other airborne equipment of our age. But explanations based on birds, hallucinations, spots in the eyes and other physiological and psychological oddities, are not strongly effected by pointing to similar experiences described in historical records.

On first sight the historical approach may appear to be free from the bias which must be expected in our contemporary communities that are under the influence of mass communications. But in present day samples such bias can be assessed to some extent if reports are studied through official agencies (Ruppelt, 1956). Obviously this is very difficult, if not impossible, for historical samples although it is possible to argue that man's attention was focussed on the heavens more strongly during some periods in history than others.

Jessup's second claim is of more relevance to his belief in intelligence from space. In his second claim it is easy to agree with statements that stones of such and such a size and weight have been transported, shaped and positioned in such and such a way,

because it can still be checked. It becomes then a question of probability estimation whether one prefers the orthodox explanation or whether one prefers help from outer space.

Of course archeological evidence supporting the orthodox explanations should be taken into account.

Let it be assumed - only for the sake of this discussion - that the observations of unidentified flying objects now and in the historical past have nothing to do with intelligence from outer space. It is hoped that this assumption is justified as an actual example, because most of the other fields of pseudo science listed by Gardner (1957) are, or could be, supported by similar historical material as was used in the field of unidentified flying objects.

Hence, even if there are reasons why this field should be taken seriously, then most likely one could select another field (e.g. from Gardner's publication) which appears utterly absurd yet for which historical support has or could be found.

Now Dodd's (1946) presentation of historical evidence from the antiquity, which may support parapsychology, is certainly a much more scholarly presentation than Jessup's, but then a good scholar could make a good case for the historical evidence of UFOs.

It seems to this author that it is difficult to find a difference in principle between the sort of historical evidence that could be presented for UFOs on the one hand and for parapsychology on the other.

When it was stated earlier that historical evidence can be used in an estimation of background probabilities, then a further qualifying procedure seems desirable. One should look at another controversial field of enquiry (e.g. UFO) which from one's own subjective considerations is likely to be wrong, and see whether the historical evidence is different in principle. If not, it provides a kind of check on one's own use of historical material in estimating the background probabilities for parapsychology.

The arguments presented here would suggest that historical materials must be viewed with much caution if it is to be used in support of a controversial field of enquiry.

There are many fields which were controversial but which can now be classified as scientifically acceptable, which were also supported by historical records. An interesting note of such historical support was provided by Kellogg (1961) when he discussed records from the antiquity, of rare and unusual behaviour of porpoises, which are now acceptable.

Murphy (1958) suggested that the history of science should be studied and in particular that periods when breakthroughs occurred, should be analysed. Kahn (1962) offered a limited amount of evidence that breakthroughs are linked with the ability to reformulate problems and that brilliant reformulations may have advanced research by generations.

It seems to this author that the spectacular leaps ahead occurred when a field was in the initial stages of investigation

with only a few active research workers. Progress is more continuous and sometimes occurs simultaneously when the field is better defined and when active research increases.

Parapsychology appears to be in a stage of development when a spectacular leap ahead is possible and certainly desirable. Yet no one can give a precise formula for a breakthrough. Kohn wrote (p.115) "Sadly enough for those inclined to practical solutions, more investigators or research money will not of itself cure this condition, but, paradoxically, serve only to exaggerate it."

This is correct if 'more investigators' follow similar attempts as have been carried out previously.

However, only a small fraction of a broad class of scientists have shown any interest or inclination to do research so far. It seems reasonable to assume that if parapsychological research could be developed on a sufficiently wide and diverse scientific basis that a wider range of problems will be formulated which should eventually lead - not necessarily in a spectacular leap but perhaps by small steps - to the same level of understanding as a breakthrough might accomplish now.

Increasing and diversifying the research basis is less ambitious than reformulating problems in a completely new way, but if the slow progress in the past is taken into account the probability seems higher that the same goal will eventually be reached.

Part I

Problems in Parapsychology

The meaning of parapsychology and methodological problems

Considering the inadequate definition of parapsychology given earlier (p. 1) it seems desirable to introduce the question of what parapsychology is, as one of the basic problems in this field.

Calling parapsychology a fringe area of psychology leaves what belongs to this area and what does not belong to it, somewhat unspecified. However the border between parapsychology and psychology should perhaps be left flexible, and while what is now called parapsychology is likely to remain, in future, a division within psychology which has its own special problems and techniques, this writer hopes that the fringe character will gradually (or perhaps suddenly) disappear.

At present one has to search for discriminating characteristics between fringe areas and the main body of psychology as well as for discriminating characteristics between the fringe area of parapsychology and other fringe areas which are not part of parapsychology.

Now, for instance, should some areas be excluded which are discussed by Gardner (1957).⁶ Obviously some fringe activities described are, at least superficially, more closely related to

6. The Fortean, p. 42 ; Lysenkoism, p. 140 ; Organon, p. 250 ; Dianetics, p. 263.

other accepted fields such as history, physics, biology etc. than to psychology. But even then it would be possible to consider their psychological aspects, and there are others which cannot be so easily identified with non-psychological fields but which should not be included under parapsychology either.

To avoid all these difficulties it might be possible simply to list the kind of events which are included in the area of parapsychology. Such a list will at least be valuable as a first introduction to this field. Rhine & Pratt (1957) provided such an introduction perhaps with special emphasis on some parapsychological events, while Murphy (1961) and Tyrell (1961) provided perhaps more balanced, if less condensed, accounts.

There is at least some disagreement (e.g. Wolstenholme & Miller (Eds.), Ciba Foundation symposium on extrasensory perception, 1956) as to the relative importance of different parapsychological events, if not about the event as such (Tyrell, 1946).

Another possibility is to consider the methods which are used in parapsychological research. This approach will no doubt introduce new problems but it is likely that a detailed discussion of specific methodological difficulties will help rather than hinder in this attempt to define parapsychological events.

It will be necessary to consider some methodological aspects in both psychology and parapsychology. On the surface it appears that there is not much difference. Indeed because of controversies and in order to exclude counter hypotheses, the methodological

standard in parapsychology is high (R. A. McConnell, 1956, p.9 ; Eysenck, 1957). Mistakes which do occur and even mistakes which could have occurred under odd circumstances, are likely to create more attention in parapsychology because there is a ready opposition waiting for an excuse to debunk the whole field (e.g. Kennedy, 1938, 1952 ; Sheffield & Kaufman, 1952 ; Pronke, 1961 ; Hansel, 1961a, 1961b). But if the question of methodology is reconsidered in a state of "dissociation" which Boring suggested in 1929 for dealing with controversies, then some agreement may be reached about principal differences and/or differences of degree.

There is a strong tendency among some contemporary parapsychologists, to emphasize the statistical basis of parapsychology. (Rhine & Pratt, 1957). But if their book is scrutinized carefully, it becomes evident that the authors do not wish to place the evidence for parapsychology on a statistical basis only. Murphy (1961, p.99) does not consider statistical evidence as a "knock-down proof" either. The continental European parapsychologists, in particular Bender and Tenhaeff, have placed more emphasis on detailed psychological investigation of single subjects, and the Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie edited by Bender has published two articles (Tornier, 1960 ; Werthmann, 1961) discussing what could probably be called the limitations of statistical evidence.

This doubt about statistical evidence must not be seen as a desire to abandon statistical procedures. At least there is no evidence of this among the authors quoted above. But to return to the question of definition and identification of parapsychology, statistical procedures should not be seen as an absolute guarantee for the existence of parapsychology but rather as a useful technique to evaluate experiments for their probable parapsychological content.

Statistical analysis is seen by these leading parapsychologists not as a tool to distinguish between parapsychological and non-parapsychological events in principle, but as a useful selection procedure.

Girdan (1962b) argued that the case for PK is "not proven" and it appears that there is no statistical "proof" for parapsychology.

The present writer objects to the term "prove" in experimental reports of standard psychological experiments as carried out by first year psychology students, because it can be argued that even in well established psychological experiments, statistical results do not provide proof of the difference between, say, an experimental and a control group, but only a probability estimation which may be expressed in a more or less confident

rejection of the null hypothesis and in more or less support (depending also on other possible counter hypotheses) of the research hypothesis to be tested.

If Girden (1962a, 1962b) was satisfied with this approach in experimental psychology, then the only justification for finding fault with this approach in parapsychology would lie in pointing towards differences between experimental psychology and parapsychology which make an essential difference in the statistical evaluation. This, however, Girden did not show.

It will be argued that such differences exist, not as principal differences, but rather as differences in degree or magnitude.

Bridgman (1956) argued against the acceptance of a parapsychological hypothesis if it is only based on a deviation from chance and not on regularities which will here be called higher order regularities. Regularities may be said to exist when manipulated changes of an independent variable(s) result in repeated observable changes of dependent variable(s).

Parapsychological experiments which do not introduce changes of an independent variable should not show higher order regularities

It remains to be seen how far this can be said of decline effects.

The problem seems to be to find a dividing line between the two classes of experiments: those that do and those that do not show higher order regularities. The extreme cases can be placed into the two classes fairly easily, but there remain experiments for which this classification is doubtful and it seems at least that some parapsychological experiments are among these.

Perhaps it is sufficient to show that some parapsychological experiments clearly belong to the class of experiments showing higher order regularities. Soal's variation of the rate of culling (Soal & Bateman, 1954) and the consequent changes in the amount of displacement, show an amount of higher order regularity which would satisfy most scientists.

However, there still exists two problems. The Soal regularities may be meaningful, (i.e. not due to chance variations), but not parapsychological in nature, i.e. they may be explained by psychological processes which belong to the main body of psychology. Hansel (Broad, 1962) made such suggestions also in connection with other key experiments (Hansel, 1961a, 1961b), which are however, generally considered to be unsatisfactory

explanations of the statistically highly significant results (Rhine & Pratt, 1961; Pratt & Woodruff, 1961; Broad, 1962).

The second problem is that since the Soal results cannot be repeated at will, the higher order regularity of a finite number of results could in principle still be due to chance. If regularity is used more broadly than suggested above, this term could describe an agreement between targets and calls as in the Soal experiments. But this is, so to speak, a first order or low order of regularity. The higher order regularity (as suggested above) comes in when additional changes occur which correspond to changes introduced into the experiment.

It could be argued that even the decline effect is a regularity in this sense. The number of trials or the sequential position of trials is the independent variable, and the amount of parapsychological response, the dependent variable. If only two values of the dependent variable (i.e. first half and second half of total result) are considered, then it is, without further numerical exemplification, obvious that such decline effects must be expected occasionally to occur by chance. It seems equally clear that if consistent higher order regularities can be found over a larger number of values of the dependent variable, then this will increase the probability considerably that such a regularity is not due to chance, but there seems to be no point between the two - value case and a "finite number of values case" where one could say that here the probability statement

of certainty.

Bridgman (1956) was correct in drawing attention to the problem of regularity, but he was not correct when he compared a parapsychological process with the event : seeing a car number plate when the probability of seeing that particular number plate is very small.

Even without higher regularities the standard ESP test is similar to predicting that a large number of particular number plates (e.g. from a particular state in the USA) will be seen during the passing of a fixed and previously specified, number of cars, and then seeing a larger number than would be expected by chance, where the probability that such a large number occurs is small.

Nevertheless this example makes it perhaps clear that it may be more meaningful to have results with a high order of regularity even if they could have occurred by chance with a probability of say, $p = 0.01$, than to have results with a low order of regularity even if p is very small in comparison.

Successful repetitions will of course increase the probability of non-chance and if carried out by different groups of experimenters, will increase even more the subjective or personal probability (Savage, 1954) estimates (of non-chance) of the readers of such reports, but again there is no sharp dividing line between the perhaps insufficient number of repetitions, (say two), and perhaps a

sufficient number of repetitions, (say 200). In a sense Soal's experiment consisted of a large number of repetitions.

The result of this argument may be summarised as follows. 7
The probability of non-chance results increases with the amount of regularity, and the subjective or personal probability estimation may even increase at a much faster rate, but there is no point (except by arbitrary definition) that could be singled out at which a probability statement would have to be accepted as certainly non-chance.

It is likely that agreement exists that in some areas of psychology experiments can be found which show less regularity than for instance Soal's experiments. Subception or subliminal perception experiments may be suitable as examples (Eriksen, 1956a, 1956b ; J. V. McConnell, Cutler & McNeil, 1958).

On the other hand it may also be taken as agreed upon that some psychological experiments have, through successful repetition, reached a degree of regularity which cannot be matched in para-psychological experimentation. This, however, would not discriminate

7. "Amount" refers to the number of changes introduced (independent variable) and measured (dependent variable) as well as the amount of agreement between these two changes.

between psychological and parapsychological experiments, but would put parapsychological results on a par with other psychological results which have not been repeated often enough, or where repetitions resulted in considerable disagreement.

The latter kind of results occur when the experimental conditions cannot be sufficiently specified. Before considering this it may be worth while to see whether statistical analysis can be of any further assistance in this attempt to clarify parapsychology.

A statistical analysis has been considered in recent years (Savage, 1954) where the acceptance of a research hypothesis is not only based on the probability at which the null hypothesis is rejected, and on the assumption that all other counter-hypotheses can be reasonably ruled out, but also on the consequences of accepting or rejecting the research hypothesis. The consequences can be of various kinds.

In a sense resistance towards the acceptance of parapsychological hypotheses may be based more on a consideration of the consequences of acceptance rather than on the probability that the null hypothesis is correct or that a non-parapsychological counter hypothesis exists.

These considerations cannot be ignored because one may expect agreement as to the calculated probability for the null hypothesis to be correct, but one cannot expect agreement at what probability the null hypothesis should be rejected (apart from

certain conventional agreements which could be arbitrarily rejected for parapsychological experiments), and one can expect even less agreement on the probability that other counter-hypotheses (e.g. fraud) exist. But such considerations will remain subjective.

It may however be possible to point, at least in principle, towards a more systematic appreciation of the consequences. This can be done by examining past experiments in psychology and parapsychology and estimating as carefully as possible how often type I and type II errors were probably committed and how often these errors were actually detected later.

Since the probability of committing type I errors is directly based on the probability at which the null hypothesis is rejected, it could be expected that in every hundred experiments in which the null hypotheses were rejected at approximately the 1 per cent level of confidence, there is on an average one type I error.

Under similar circumstances the probability of committing type I errors will again be quite similar in psychological and parapsychological experiments, but this may not be the case for detecting type I errors.

The remainder of the argument has to be speculative by necessity but it does not seem unreasonable to assume that the probability of detecting type I errors in a parapsychological experiment is lower than in psychological experiments. Detecting

a type I error in parapsychology means after all, finding some good reasons why the results of an experiment which were previously interpreted to be of parapsychological nature, should be interpreted as being due to chance.

Probably all serious parapsychologists are aware that the result of a significance test supporting a parapsychological hypothesis is only a probability statement, but there seems to be no way of deciding that one particular experiment gave 'Schein-signifikanzen' (significant chance results) rather than another. It has been argued (Rhine & Pratt, 1957 ; Rhine, 1962) that some experimenters do not have the ability (in terms of unspecified personality characteristics, i.e. distinct from technical know-how of procedure etc.) to carry out successful parapsychological experiments and that the ability to do so may fluctuate or be lost altogether.

One could perhaps assume that an experimenter who was unsuccessful in 20 experiments and successful in one, based his success on a type I error.

It may be argued that in comparison more definite reasons may be found in the field of psychology to suspect the occurrence of Type I errors occasionally and one may further speculate that type I errors may actually be found more often in psychology than in parapsychology although even in psychology the detection is likely to be limited to a small percentage of the actual number

of type I errors which one must expect on theoretical grounds.

This speculation could be tested to some extent by a well planned and organized and rather gigantic survey covering a large area of psychological and parapsychological experimental work carried out say during the last five years, to find out how often type I errors were ever recognised as such. If such a survey were undertaken on a sufficiently large scale it could lead to useful information showing not only differences (if any) between parapsychological and psychological experiments, but also between different areas and types of experiments within the general field of psychology.

Type II errors could be investigated in a similar way but since the probability of committing type II errors depends also on the number of measures taken into consideration for analysis these results of a survey may be more difficult to classify.

If differences could be found and to some extent be evaluated it would be possible to suggest different significance levels as a guide for further evaluation. However from a certain point onwards the consequences will have to be considered in non-statistical terms, but numerical results (from the suggested survey) may be helpful in this respect.

If it should be found for instance, that on the basis of similar expectancies, type I errors are detected in psychology three times as often as in parapsychology, then it might be reasonable to suggest a significance level for parapsychology at which the necessary probability is three times smaller than in experimental

psychology. This however would also increase (even if not to the same extent) the probability of committing type II errors. Because of this, non-statistical considerations such as harmful effects through application or publication (if type I errors occur) i.e. wrong diagnosis from psychological tests and wrong channelling of effort in either psychological or parapsychological research, must be contrasted with, and weighed against, other harmful research effects such as abandoning an important line of investigation (possibly because no further finances were available after official results were labelled "non-significant") or non-publication of important results if type II errors occur.

Guilford (1956, p. 216), Soal & Bateman (1954, p. 42), and Bender (1960), have argued for higher significance levels in parapsychology, and parapsychological experiments which were introduced as "key" experiments (Soal, 1954 ; Mangan, 1959 ; Broad, 1962) usually had very high probabilities at which the null hypotheses were rejected. On the other hand, very high 'Scheinsignifikanzen' have also been found when lists of numbers were examined which are believed to be random numbers (Oram, 1954 ; Werthmann, 1961).

It may be argued that an analysis of the occurrence of type I and type II errors in psychology and parapsychology may lead to quantifiable differences between the two fields. The meaning of such differences, i.e. the recommended interpretation of results,

must however, also be based on non-statistical considerations.

This statistical discussion has so far shown no principal differences between parapsychological and psychological experiments. A probable difference of degree was noticed between the amount of regularity and a possible difference of degree in the detection of type I and type II errors was suggested.

It may also be argued that in most psychological experiments the ratio of known to unknown variables is more favourable than in parapsychological experiments. It is possible to suggest further, that statistical procedures have only been tested successfully under the more favourable conditions and that no a priori justification exists to apply the same statistical procedure with equal validity expectations to the less favourable conditions. Tornier (1960, 1963) formulated this more precisely by saying that in the case of the favourable ratio the effect to be demonstrated and identified on the basis of significant tests will show up "fast" compared with possible effects which may be "in" a random number table. Empirical experiments have shown that under these "fast" conditions, significance tests are justified but this becomes increasingly doubtful as the effect being measured "slows down". In a parapsychological test situation Tornier (1963) would argue that if only a small proportion of responses suggest parapsychological results, that this is a more doubtful situation for statistical analysis compared to one where the proportion is higher.

It is perhaps possible to find clear differences of this sort between some psychological experiments and parapsychological experiments in general, but one may also find psychological experiments where the known to unknown ratio is similar to the parapsychological situation. (p. 35).

Statistical analysis does not seem to lead to a unique identification of parapsychological events. But once the limitation of statistical applications are appreciated the quantitative method in parapsychology will be seen as a highly desirable one and not as fundamentally different from methods used in psychology. To define parapsychology as a fringe area of psychology seems supported by these considerations.

It was argued initially (p. 39) that a clear distinction between the field of psychology and the field of parapsychology cannot be easily established. Discussing statistical procedures has not lead to such a clear distinction, but once it is understood to what degree methods in experimental psychology are essentially suitable for parapsychology research, the methodological approach may be used to discriminate between parapsychology and other fringe activities; as for instance, discussed by Gardner (1957). Dianetics (Chapter 22) could be ruled out because parapsychological and psychological methodology is not suitable.

Parapsychological methodology is sufficiently clear to allow serious opponents to pinpoint specific statements with which they disagree and allows the supporters of parapsychology to reply to such specific criticism. (Hansel, 1961a, 1961b ; Rhine & Pratt, 1961; Pratt & Woodruff, 1961).

How far disagreement can be resolved is partly a measure of how scientific (if at all) parapsychology can be said to be. There seems to be some disagreement about the meaning of "scientific" and because of this, disagreement among parapsychologists, as to whether parapsychological hypotheses may be called scientific. Rhine & Pratt (1957) clearly introduce parapsychology as a science, while Murphy's (1961, p.1) statement seems much more conservative.

"Psychical research, or parapsychology, consists of observations recorded in a form which aims at order and intelligibility, but which cannot by any stretch of the imagination be subsumed under the science of today."

The difference may be less real than it appears. Rhine & Pratt are mainly concerned with methods (which they postulate should be scientific) while Murphy's remarks are possibly based on the difficulties of replication in parapsychology.

Nevertheless a discussion of the scientific aspects of parapsychology may help to clarify the subject matter further.

McQuigan (1960) defined science as "the application of scientific method to meaningful problems" (p. 3). Scientific method cannot be defined easily, if at all, in a short sentence, nevertheless it can be described quite adequately by giving principle

examples of the kind of experimental procedures which make up scientific method. This amounts to an operational definition of the method which unfortunately is rather lengthy. Scientific method through principal examples has not only been described adequately in psychology (McGuigan, 1960 ; Scriven, 1961) but also in parapsychology (Soal & Bateman, 1954 ; Rhine & Pratt, 1957 ; Tyrrell, 1961 ; Murphy, 1961). But there is still the meaning of "meaningful problem" left. McGuigan argued that meaningful problems can be expressed in hypotheses whose propositions can be determined to have degrees of probability. If such probabilities cannot be established at present but may be established in future (e.g. after inventing some suitable testing apparatus), then the problem has only potential meaning.

It seems then that this definition would group single parapsychological experiments into the class of scientific experiments, at least in as far as one is satisfied with probability estimations in parapsychology. But the situation is more difficult with respect to the total area of parapsychology or with respect to branches such as ESP, PK, etc. It would be exceedingly difficult to find a degree of probability in any of these branches and consequently for parapsychology as a whole, and Rhine (1956) has argued previously (in connection with the Price controversy) that no crucial test (which would determine such a degree of probability) can be suggested at present. However, if the situation is compared with psychology, the

difference is less favourable for psychology than one might hope. Many clinical and psychoanalytical techniques have only limited use of scientific method and have certainly no crucial experiments, but even such a rigorous approach as Hull's (1952) can lead to conditions for which a degree of probability cannot be determined as was indirectly pointed out by Broadbent (1958, p.9).

Of course not all scientists would classify psychology as a science (Sullivan, 1949 ; Hoyle, 1957) and most would agree that not all psychological work is scientific.

Scientific method then seems to be useful and desirable in parapsychology but there is still some doubt how far parapsychology can be called a science, but it is also doubtful whether psychology can be said to be entirely scientific.

But if a probability can be found for a single experiment as to its parapsychological content then it ought to be possible to estimate a probability for the total number of experiments in the total field or in any particular section. In a limited way this was done in the early stages of organised quantitative research (Rhine, et al, 1940).

This leads to one of the current practical problems in parapsychology. Rhine (1956, 1959) has argued that no crucial experiments for the parapsychological hypothesis can be set up but also that each experiment must be considered a unique effort to show evidence of parapsychological events. Or in other words,

no two experiments are really the same. Most psychologists would agree with this but some might argue that an event can only be demonstrated if the experimental situation can be sufficiently specified such that changes from one experiment to the next will not influence the results to be demonstrated in a significant degree. Such repetitions have been demonstrated in parapsychology (e.g. Soal & Beteman, 1954 ; van Busschbach, 1956) and discussed (e.g. Wolstenholme & Miller (Eds.) 1956 ; Murphy, 1958, 1962).

Nevertheless there is also a good deal of evidence for drastic changes without sufficient knowledge to explain them in any satisfactory way (Eisenbud et al, 1960, p.228 ; Murphy, 1961).

The historical events leading to Soal's and Thouless' discovery of parapsychological results, and Pratt's example and remarks (1942, p. 136) "Even experiments made on the assumption that ESP is impossible have in some instances produced evidence to the contrary:" suggest that it may be equally difficult to set up psychological conditions under which subjects will show no parapsychological evidence. This has not been tested sufficiently. If psychological conditions can be found under which no parapsychological results occur - although this is equally difficult to demonstrate - then this could be of interest for research in this field.

On the other hand Schneidler's (1945, 1949) and Humphrey's (1949, 1950) experiments seem to suggest that there are some

means by which the amount of parapsychological output can be predicted, and consequently it does not seem impossible, in principle, to set up experiments which turn out chance results and nothing else. Rhine (1936, 1939) would say that such tests are not parapsychological tests in the first place and that these results do not in any way effect other results which are supposed to be parapsychological in nature.

While there are some ways which improve the likelihood of obtaining parapsychological results (e.g. Schmeidler's sheep, goat method, other personality measurements, and/or working with children), Rhine cannot specify a set of conditions (including any experiment he likes to choose) which will definitely provide him with parapsychological results. Because of this it is not really possible to say that in the experiments where significant results were obtained, the right parapsychological conditions existed and in the experiments where no significant results were obtained, conditions existed which were unfavourable to the demonstration of parapsychological events.

The possibility of an accumulation of type I errors cannot be ruled out unless either single parapsychological experiments are evaluated on the basis of other non-significant experiments (carried out by anybody) or if experimental conditions can be specified to such an extent that other experiments can be clearly distinguished on the basis of such specifications.

Gardner (1957) argued against parapsychological results

by assuming an accumulation of type I errors and by referring to the use of several possible and independent hypotheses which are tested simultaneously, and where any one can demonstrate parapsychological events. For example, correct targets may not reach a significant deviation from chance in a PK experiment, but a significant decline effect may be taken to demonstrate parapsychological events (R. A. McConnell, Snowden, & Powell, 1955).

However the procedure for dealing with alternative hypotheses is fairly simple and has been followed in parapsychological experimentation. If two hypotheses are tested simultaneously where each one may demonstrate parapsychological events alone, then if only one is significant, this significance must be twice as high as if only one parapsychological hypothesis is tested.

If one is hostile to parapsychology it is possible to speculate that the simultaneous hypotheses have introduced an increased possibility of self-deception, but there is no doubt that the problem of alternative hypothesis testing is recognised and adequately solved if necessary (Rhine & Pratt, 1957, p. 171).

Rhine (1959) discussed what he called the pooling fallacy. In presenting the view of those writers who adhere to this fallacy Rhine said (p. 39) "The implication is that no unit of work relying on statistics, no matter how independently conducted or how well designed, can be considered entirely by itself - on its own merits."

It seems that parapsychologists agree with this view when 'unit of work' in the quotation is replaced by 'whole field of say PK in parapsychology' (p. 29).

The quotation as it stands seems to this author logically correct. How far the consequences, that is taking into account all other experiments in the statistical analysis, should be followed, is a question of practical convenience. It does not follow as Rhine seems to argue, that because such a demand is impracticable, that it is also a fallacy.

"How could a research worker ever know where he stood until all reports were in? And how could unpublished work be rounded up? Probably no-one would insist on carrying pooling so far, but this absurd extreme only serves to point out the fallacy that runs through the whole idea of compulsory pooling." (Rhine, p.39).

Rhine was quite correct when he pointed out in the same publication that if the demand for pooling is accepted then in other fields of experimentation, e.g. experimental psychology, pooling would also have to be introduced.

This somewhat frightening structure of logical consequences can be brought down to a level of practical convenience if the following questions are asked:

1. How many similar experiments are likely to be carried out which are published strictly in accordance with the statistical significance of results? (i.e. publication

if significance is reached, no publication if no significance is reached).

2. Has an independent variable been tested over a wide range or in other words has a high order of regularity been established?

Pooling is of little practical necessity in most experimental fields because the numbers given as answers to question 1 are usually low and to question 2 usually high compared to parapsychology. It is because of this kind of difference between the two fields of enquiry that pooling may be a problem in parapsychology while generally it is not a problem in other experimental fields.

However if conditions exist in any experimental field which makes a high figure (for question one) and a low figure (for question two) likely, and particularly if the investigated events cannot be repeated on demand, the problem of pooling should be taken seriously or in other words, the question should be considered whether statistically significant results should be expected to have occurred by chance because of an accumulation of ^{type} ~~ne~~ I errors.

The problem of accumulating type I errors is not easily solved. This can perhaps be illustrated by using some simple theoretical examples.

1. If 10 independent persons each have a different system and if they each select a number (according to their

system) from a different lottery (where each lottery has a winning probability of $P_w = 0.01$) then there is a probability of $p = 0.1$ that one of these 10 persons will select a winning number. If one does select a winning number he will probably argue that he had only a 0.01 chance of winning and he might become very confident about his system.

2. If one person develops 10 independent systems and tries them out by selecting one number from 10 different lotteries (each according to one system), then there exists again a probability of 0.1 of drawing a winning number (P_w as in the previous example).

The examples do not change if instead of 10 different lotteries only one is used.

It seems that Rhine's single experiment (1959) corresponds very closely to the first example and that the statistical significance of any one experiment can only be estimated if all other experiments are known.

If nine attempts to obtain parapsychological results are never published then the tenth attempt can only be said to be significant if its significance is ten times as high as requested for a single attempt. Unfortunately it does not seem to matter who the different people are or whether their experiments are precisely the same, as long as no difference between parapsychological

test conditions and non-parapsychological conditions can be specified. If the probability is high that any one of the ten will publish the results if successful, but not if unsuccessful, an accumulation of type I errors is possible.

Of course it is also possible, and on subjective estimation even probable, that at least a large number of the kind of experiments which only superficially resemble parapsychological experiments - (e.g. Sheffield, Kaufmann, 1952) - and which Rhine wants to have excluded - are in fact substantially different. But the difficulty is that on the one hand at present no discriminating specifications can be given, and on the other hand Parapsychologists reject few (if any) positive results as probable type I errors, or as due to other counter-hypotheses even when these results were arrived at under conditions which could have led to their possible exclusion from the field of parapsychology, if the null hypothesis had to be accepted. Yet among hostile and careless experimentation which is perhaps not so likely to demonstrate parapsychological events, some type I errors must be expected.

The argument against excluding any insignificant results works both ways. Both Murphy (1961) and Tyrrell (1961) mentioned significant results which were not published because the experimenters were hostile to parapsychology and expected

insignificant results, or because the experimenters did not get the kind of significance results they expected on some theoretical grounds. The present writer knows of at least one such case. It is therefore by no means certain whether Gardner's (1957) suspicion that the significant results reported in the Journals are actually type I errors, would be supported if all unpublished parapsychological works were made available.

To carry the example No. 1 somewhat further, it is possible, depending on how the lottery is drawn, that the "system" used by the winning person is in fact a method which really increased the probability of selecting the correct number (a mechanical number selector may jam at certain positions and from previous results such positions and corresponding numbers may be estimated). If this is the case then the 'system' should be successful in further number selections. If the second choice of a number is correct again, what probabilities are involved? Since only the winner of the lottery in the first round is selected as the winner for the lottery in the second round, and since any other persons winning in this second round are discarded as meaningless in the investigation of "successful systems", the probability of the one selected person winning in the second round is 0.01 and independent of the number of other persons trying again.

The probability of winning a second time in this example is $0.1 \times 0.01 = 0.001$.

Unfortunately this example is somewhat over-simplified because it assumes that all first round winners are known and that in fact only one winner emerged from this first round.

If N lotteries are envisaged which together offer n numbers to be selected with $p = 0.01$ to win, then (after all numbers have been selected), it must be expected that $0.01 \times n$ persons will win by chance (if one person is restricted to one number).

If only one person is selected to win again, then his probability to win again is independent of other participants but if a winner in the second round is predicted only on the basis of having won previously, and if the number of such previous winners who enter the second round is not known, then it is still possible that a particular double win may occur within a probability which is really too large to reject the null hypothesis. Yet if the number of first round winners entering the second round is either unknown or ignored, the null hypothesis is likely to be rejected by the investigator.

Nevertheless it seems clear that the danger of accumulating type I errors is reduced substantially with only one successful repetition.

The first example seems to indicate that there is some danger of an accumulation of type I errors if parapsychological test conditions cannot be sufficiently specified. Attempts to specify such conditions should perhaps be made in spite of the

danger that such specifications may lead to an unsuccessful demonstration experiment (Rhine, 1956).

The example also shows that at least one repetition of an experiment is highly desirable as this would decrease the possibility of type I errors considerably.

This discussion is also relevant to experiments in psychology, particularly if external criteria which support the statistical findings are fairly unreliable. The second example (p. 51) does not add much to this argument but it may be used to illustrate the possibility of unintentional self-deception in psychology as well as parapsychology. The practical consequences of rejecting certain experiments as non-parapsychological experiments in statistical evaluations, are probably not important in the sense that the overall picture of parapsychological evidence is not likely to change much (Whittlesey, Greenwood, 1959, p. 135). This is partly due to the very high significance of some test results (e.g. Pratt & Woodruff, 1939 ; Soal & Bateman, 1954 ; Rhine & Pratt, 1954) and to the actual repetition (even if limited in number) of some experiments (e.g. Deguine, 1959), but also due to the probability that among the totality of unpublished results, significant results (whether due to type I error or not is of no importance in this connection) will be found.

Perhaps it can be said that experimental conditions can be specified somewhat better than has been suggested so far.

In particular the successful experimenter is singled out as one of the test conditions which are essential for parapsychological events (Rhine, 1939). It was, however, already noted earlier (p. ⁴⁹) that success does not clearly depend on the original expectations and hopes of the experimenter nor does initial success in any way guarantee continued success (Dale & Woodruff, 1947 ; Van Busschbach, 1961 ; Ryzl, 1963). ⁸ Schmiedler and McConnell's results (1958, p.46) also indicate that the attitude toward parapsychology (as tested in a certain arbitrary and probably rather limited way) does not lead to clear results in individual cases. That is although group results differed significantly in a way that might be expected from the expressed attitudes, individual results ran counter in both groups. This again might be due to the limitations of the "attitude tests" which Schmiedler used, but if compared with Humphrey's (1951, p. 461) remarks ⁹ it is

8. See also Ceover's experiment at Stanford in Parapsychol. Bull., 1962, No. 63.

9. "The trend shown in the early series has not held up consistently. In fairness to the measure it should perhaps be mentioned that the later series include widely differing conditions; and it is almost certain that the effectiveness of any measure of personality is the function of the total experimental atmosphere, including experimenter's personality, instructions to subjects, type of ESP test, physical environment, etc."

probably fair to say that there is as yet no definite way of selecting single successful or unsuccessful subjects, and that there is as yet no way of knowing whether subjects who are initially successful will stay that way and for what length of time. Because of this it seems highly doubtful whether any set of definite criteria exist which can discriminate between those experimenters who have the "ability" to conduct parapsychological experiments and those who do not have this ability. (As on p. 38 ability refers to unspecified personality characteristics and not to technical know-how.")

The problem of repeatability in parapsychology is generally recognised. Murphy (1961) linked it with the term scientific and suggested that experiments can be called scientific only if they can be repeated again and again.

Rhine (1959) argued that in some recognised branches of science e.g. geology, repeatability is generally impossible. But even in physics some events cannot be repeated on command (R.A. McConnell, 1947).

Schmeidler (1959) spoke against the direct repetition of experiments since they are not likely to lead to new knowledge. This is quite correct. On the other hand even a limited repetition helps to minimize the possible occurrence of Type 1 errors. (p. 54). A practical suggestion may be to define some aspects of repetition more precisely (see also Wolstenholme & Miller, 1956, pp. 39-49).

Obviously an exact repetition is impossible anyway and the question is what should be accepted as an attempt to repeat an experiment and what should be regarded as a new and different experiment.

On the basis of the discussion on the pooling problem (pp. 49-50) it can be said that the danger of accumulating type I errors could be equally high if only a few experiments are (by themselves) significant out of a large number of experiments which are intentionally different, compared with a large number of experiments which are as similar as possible.

Since a statistically insignificant experiment tends to increase the possibility of Type I errors for a different significance experiment, two - in some aspects intentionally different - experiments which are both significant could support each other.

Hence the limitation of the term repetition is arbitrary. While it is undesirable to claim support from an experiment which appears to be quite different, an intentional alteration of some aspects of a previous experiment does not exclude the applicability of the term repetition.

It seems that an appreciation of the pooling problem will be a help rather than a hinderance in the repetition of different experiments. Nevertheless it is also desirable to specify in advance what difference in the results are expected because of intentional differences in those experiments which may count as repetitions.

If it is possible to predict which particular experiments out of a large number of experiments will produce significant results and which experiments will only show insignificant results, then even the insignificant results help to support the total experimental evidence at hand.

Without any prediction there exists some danger that after a successful experiment (E1), a second experiment (E2) is regarded as a repetition if successful, but as a different (non-repetition) experiment if unsuccessful (Murphy, 1959, p.133). Moreover without a prediction it seems necessary to pool E1 and E2 unless "repetition" is used in a very restricted sense. But then it is doubtful whether any parapsychological experiment has ever been repeated and this view is not adopted among parapsychologists.

It is generally argued that at present experiments in parapsychology cannot be repeated at will (Rhine, 1959). But it has also been stated that successful repetitions have been carried out (Soal & Bateman, 1934 ; Murphy, 1961, 1962).

The ESP tests in schools with children have resulted in a number of successful experiments some of which were intended primarily as repetitions of previous ones. (van Busschbach, 1956 ; Deguise, 1959)

Ryzl's experiments (Ryzl & Ryzlova, 1962 ; Ryzl & Pratt, 1962, 1963a, 1963b, 1963c) can also be regarded as repetitions.

The experiments with children and Ryzl's use of hypnosis are

relevant to the experiments reported here because in both cases the specific parapsychological experimenter-subject relationship was changed into more standardisable relationships : teacher-pupil relationship and experimenter inducing hypnosis-subject relationship.

The subjects in these experiments were not as much (if at all) involved in parapsychological problems and any enthusiasm generated by the experimenters was not specifically concerned with parapsychological issues.

It seems to this author an important attempt to depart from the experimenter-subject relationships which are based on converting subjects to be interested in or devoted to parapsychological task which is often boring (e.g. calling ESP cards) because the kind of initial necessary attitude of the experimenter is difficult to maintain, more difficult to describe and unlikely to be brought under experimental control for some considerable time.

Experiments with children have not continued to be successful in every case (e.g. van Busschbach, 1961) and the stability in Ryzl's experiments is also limited (Parapsychol. Bull., 1963, No. 67). Nevertheless the two kinds of experiments mentioned above were among the more important contributions in the last decade, and while it is impossible to say with any certainty whether the shift away from parapsychologically centred experiments had anything to do with this success, to look at the above research in this way may provide the basis for a reformulation of problems which Kohn (1962)

found important in an analysis of scientific break-throughs.

It would, however, be misleading to regard the above situations which were used experimentally, as simple. They are, as far as can be judged now, highly complex, yet the complexity can perhaps be handled with more hope of success. It does not appear a hopeless task to define some necessary response level to, say, a motion picture of a fairy tale presented to a group of children, at which they ^{can} be regarded as sufficiently involved in the story to participate at some stage parapsychologically. Perhaps a successful experiment will never be accomplished under such conditions but it seems the sort of thing that ought to be tried out in order to improve the repeatability of an experiment.

If this sort of approach works once it can be expected that it has a reasonable probability not only to work again with the same parapsychologist as experimenter (now far removed into the background) but one would hope that any interested scientist has a similar probability to succeed.

What makes parapsychology less acceptable to scientists is not so much the prospect of getting some experiments which show no significant results, but the almost magic distinction between a successful and an unsuccessful experimenter in parapsychology (Rhine, 1959), with little concrete advice as to what personality factors or other peculiarities are involved.

In an attempt to take some initial steps towards the solution

of this problem the author has carried out a number of disguised experiments with adults as subjects (pp. 123,150,208). In an undisguised test favourite music targets were used in an attempt to provide a strong and in a sense automatic, (i.e. to some extent independent of the experimenter and test conditions) "personal involvement variable" to see whether a parapsychological test can be stabilized under such conditions.

Emotions and 'surplus' meaning in science and parapsychology

There are a number of remarks in parapsychological literature (Woodruff, 1961, 1962) directed against critical discussions of parapsychological results by professional people who are not active parapsychological experimenters. This, however, provides a situation for, as an example, the sceptical scientist (who is nevertheless interested, and this is a very strong point in his favour) which is perhaps not quite fair, because to discuss parapsychology critically and adequately he is first asked to achieve significant results in a parapsychological experiment, for without significant results his research is not considered parapsychological (Rhine, 1939).

In a sense he is asked to admit parapsychological results before he is allowed to discuss them critically. It is doubtful whether it is fully justifiable to demand experimental participation in parapsychology of the trained critic. Woodruff's example of counting the teeth in the horse's mouth carries a certain amount of emotional overtone which will be discussed later (p. 21).

The important point is that a trained scientist should be able to judge from the experimental report what was done in the experiment and if this is not possible from the report, then

it is essentially a shortcoming of the report. It is possible to refer to terms such as "the atmosphere of an experiment" (p.⁵⁶) which is supposed to express certain characteristics of the experiment which cannot be defined in a report and which is best experienced through experimentation. But once any such indefinable experimental characteristics are admitted, the experimenter could be accused of being too involved and the critical scientist would perhaps be justified in staying away.

More participation (even if only through critical discussions) and interest in parapsychology by professional scientists, seems to be largely desirable at this stage. Encouragement towards criticism may lead to unreasonable, uninformed and time-wasting discussions, but to this writer such dangers seem to be small compared to discouraging useful critics, for instance Scriven who with Meehl (Meehl & Scriven, 1956) ably defended parapsychology in the G. Price controversy (1955) and compared to discouraging experimenters on highly doubtful assumptions because they were initially unsuccessful.

The question of criticism is closely linked with the problem of emotional reporting. Emotional reports on scientific or pseudo-scientific matters can be highly entertaining as the quotations in Gardner's book (1957) indicate. Unemotional writing can be rather dull although it may be occasionally unintentionally amusing as in the passage by Asher (1958, p. 502)

quoted by McGuigan (1960, p.77) "Overconscientious anonymity" can be overdone as in the article by two authors which had the footnote "Since this article was written, unfortunately one of us has died."

To add emotional toning to reports or criticism of reports can, apart from entertainment, have the following functions. It may help the author to regain a desirable state of emotional balance. It may do the same for some of the majority of the readers, but it will also as a rule, tend to add support to reported events, procedures and to the analysis or criticism of such events and procedures. It may also help to establish group behaviour or in a wider sense, for instance, schools of psychology.

Although emotional writing whether intentional or unintentional, is supposed to support the issue at hand, it seems highly doubtful whether anything like that is really achieved because after all if the case is convincing enough, emotional support is hardly necessary and one may suspect that emotions are added if the case is not too clear on methodological grounds. Once an argument has led to certain results it may be legitimate to illustrate these and this will often lead to unavoidable distortions, but the right to illustrate is not the right to illustrate emotionally. Pratt (1963) used an emotional illustration when he "placed the Girden-Murphy controversy (Girden, 1962a; Murphy, 1962) into proper perspective":

"In doing this he (Murphy) pointed out Dr. Girden's shortcomings with the happy combination of gentleness and firmness that the wise teacher might use in correcting a student who has misbehaved in the expectation of being secretly applauded by his classmates" (Pratt, 1963, p. 199).

There is a good deal of evidence in Girden's paper which is not restricted to a sober discussion of the work at hand. This was noted by Murphy (1962). Girden (1962a) quoted Soal against Rhine with more than necessary emphasis (p. 382) but it seems highly doubtful whether emotions are desirable even if they are used in replies to emotional statements. Murphy's report was remarkably free of such emotional overtones but it seems to this writer that Pratt's (1963) was not. Moreover Pratt included items of information which, although not emotionally toned themselves, are likely to create emotional bias among readers.

It may perhaps still be fair enough to inform the reader that Girden spent little time at the Duke laboratory for parapsychology and that he did not return after a short visit, but it seems quite irrelevant to the discussion at hand that Girden did not attend a talk by Rhine at Brooklyn College (Pratt, p. 201). Is the reader of Pratt's article expected to assume that Girden was not really interested in obtaining any information about PK? If this was Pratt's intention then the reader is entitled to a good deal more additional informatio

as without this such an assumption would be unjustified on the basis of the statement that Girden did not attend Rhine's talk. If it was not Pratt's intention then what was the reason for including such a statement? Although this recent exchange was fairly mild in emotional overtones, it is unfortunately not the only one, nor is it restricted to parapsychology. Boring (1929) noted that if emotions are expressed in a controversy, they are likely to increase in intensity as the exchange continues and that they do not help in sorting out the problems at hand.

Even in physics emotional outbursts can interfere with the evaluation of arguments. Bridgman (1958) noted this when he discussed indeterminism (pp. 50-51 and 55-56).

"It seems to me that most of the criticism of Bohr arise from unwillingness to accept the conditions of the problem Bohr has effectively set himself. Perhaps part of the unwillingness arises from failure to appreciate the purely formal aspects of the problem of devising methods of thinking about a hypothetical physical situation. Regarding the problem in this formal light, one should be able to attack it purely as an intellectual exercise. But I believe the unwillingness actually involves many other factors, some of them recognizably emotional. The reason people will not see the problem in a purely formal light is that the hypothetical state of affairs basic to the problem is so

much at variance with their conventional and traditional pictures of what the experimental situation must be that they will have none of it, and even refuse to speculate how they might act if it existed. The repugnance of different persons to accept the possibility that the world is actually constructed according to the hypothesis of orthodox quantum theory varies greatly and may involve considerations closely approaching the religious. This is perhaps most strikingly shown by Einstein, who could not bring himself to accept the idea that chance plays a fundamental role in the scheme of things, and who passionately exclaimed, "der Herr Gott würfelt nicht" ("The Lord God does not throw dice"). Einstein's repugnance led him so far that, instead of postulating that there might be experimental facts not yet discovered, which is perfectly tenable position and all that he needed, he was convinced that no theory giving a fundamental place to probability could be logically consistent, and he spent a great deal of time trying to point out logically untenable aspects of quantum theory, always to be patiently refuted by Bohr (pp. 50-51).

I find it hard, in reading the recent discussion of causality and determinism, to resist the impression that many of the debaters were influenced by extrascientific considerations. This influence is evident even in the work of Max Born who says, in combating Schrödinger's thesis that it is waves and not

particles that are fundamental, "I think Schrödinger's suggestion is impracticable and against the spirit of the time." This is perhaps unfair to Born, who may have been throwing back into Schrödinger's teeth his demand that the theorist be aware of cultural background, but nevertheless one could wish that Born had not said it.

Among writers of less scientific stature than Born the influence of extrascientific factors is unmistakable. This is particularly evident in Mario Bunge's paper "Strife About Complementarity." The burden of Bunge's paper is that physicists are at last awakening from the "dogmatic slumber" in which they have accepted "the official philosophy of quantum theory, which is essentially of a positivistic character" and are embarking instead on "new realistic, rationalistic, and deterministic trends." It seems to me that there is too little argument in the paper and too much name-calling. It is assumed that the reader will react negatively to such epithets as "positivistic" and "empiristic" and positively to "realistic", "deterministic", and "scientific materialism." The assumption that the reader will react in the expected way to these epithets and that it is desirable that he should so react obviously does not have its origin in any purely scientific experience." (pp. 55-56).

It seems that Bridgman was particularly concerned about his own side, so to speak; this seems to be a desirable procedure which has been adopted in this thesis.

Rhine & Pratt (1961) in a reply to Hansel's paper (1961a) pointed out that the scientific status of parapsychology depends on the stability of this field when under criticism. It seems to this writer that Hansel's attack on the Pearce-Pratt experiment was particularly weak and entirely rebutted by Rhine & Pratt on methodological grounds. It seems however, that the scientific status of the rebuttal was weakened rather than strengthened when information was included which was not relevant to the issue.

Even indignation about the charge of fraud seems out of place here because Rhine & Pratt agree in the same paper that parapsychological experiments should aim to exclude the counter hypothesis of fraud and that this was successfully done in the advanced Pearce-Pratt series.

If fraud is an accepted point of critical discussion, then Hansel's papers were indeed temperate as was pointed out by Scriven elsewhere (1961). The irrelevant information supplied by Rhine & Pratt is that which refers to the finance supplied by the Duke Laboratory (which had already been acknowledged by Hansel) and that referring to Hansel's unwillingness to discuss the questions.

It may very well have been discourteous of Hansel to refuse discussions but once his papers were accepted for publication

by the Journal of Parapsychology, the only replies which in conformity with accepted scientific practice should have been offered, are those parts of the rebuttal which are concerned with the actual contents of Hansel's papers.

It is precisely through the kind of introductory remarks - which in the days of limited scientific journal space should have been returned for revision by the editors - which give outside scientists the impression that the field of parapsychology is still far removed from a science. This is particularly unfortunate because in this controversy the rebuttal of Hansel's criticism about the Pearce-Pratt series seems quite sound on methodological grounds.

No final evaluation of the replies by Pratt & Woodruff (1961) to Hansel's criticism is given here since a more detailed study of material which is not readily available, would be necessary. It seems nevertheless likely that in this case also Hansel's criticism is unjustified.

In another case of controversy, this time between members of the Parapsychological Association (Scriven, 1961, 1962a, 1962b ; Woodruff, 1961, 1962 ; Stevenson, 1962 ; Onetto, 1962) perhaps the important issues raised in the original paper by Scriven were obscured by more or less emotional replies. Scriven's original

contribution suffered from some lack of information about recent publications, but the content of his paper was presented without undue emotionality and this can also be said about his rejoinders although one may have a slight uneasiness about such expressions as "goodbye physics", (1961, p. 312) "eat the cake and have it" (1962b, p. 132). The replies to Scriven's remarks unfortunately show a good deal of emotionality, e.g. "failure of Dr. Scriven to complete his homework properly", Francis Bacon's story of the number of teeth in a horse's mouth from Munn's textbook, "descent from the self created Olympian heights" - (Woodruff, 1961, pp. 267-268) and "assume the posture of a tough minded critic", "homework" - (Stevenson, 1962, p.64).

The question that must be asked again is, do these emotional replies indicate scientific status, and one feels uneasy about the answers.

It seems however that the most important point in Scriven's paper was only taken up by Stevenson. This is the problem of spontaneous case studies in parapsychology. Scriven interpreted research findings by Walter (1960) as an indication that one must have doubts in one's subjective estimation as to whether any particular case has parapsychological content or not. This seems to be entirely justified on the basis of Walter's statement (p.22).

"Applying these observations and conjectures to the general problem of illusory experiences, it would seem worth considering whether the liability to "spontaneous" experience

is related to these same features of brain activity. The neuronic complexity of the brain is so vast and its metabolic economy so intricate that conditions similar to those contrived in the laboratory must arise quite often by chance in everyone. The proposition is that in certain people such events would set up states of activation not very unlike those of everyday mental life - perhaps more vivid or less obviously relevant, but not qualitatively unfamiliar. But in others - actually the majority of the population as a whole - such states would be rare and inexplicable experiences, comparable with hearing a familiar and urgent voice speaking a strange language. The subjective interpretation of such experiences would obviously differ, on this hypothesis, according to the previous habitual mode of imagination. At one extreme the habitual visualist would dismiss his occasional illusions as trivial exaggerations of his normal state and would maintain a stout scepticism as to their interest or relevance to outside events. At the other, the person unfamiliar with illusory sensations, or images as a part of his mental machinery would regard them as supernatural manifestations, related not to his own internal predicament but to some external phenomenon."

Obviously this is of outstanding importance as it involves a basic issue in parapsychology. Whether terms like 'absorption' and 'evaporation' (Scriven, 1961) are justified, is in comparison less crucial. Scriven did not claim that there are no case results which could have had and will have parapsychological content, but he does conclude from Walter's finding that it is now extremely difficult to estimate any probabilities for the parapsychological contents of such cases and that previous estimations (Tyrell, 1961, pp.17-48 ; Broad, 1962, pp.99-112) must be considered to be too favourable for parapsychology. However Scriven (1961) emphasised the difficulties in obtaining results which could throw some light on this problem (e.g. through a more sophisticated census) more than Walter (1960, p.23). And this discussion does not reduce the possible use of spontaneous case studies for further quantitative research. There are and there will be cases left where one's subjective probability estimation will still favour a parapsychological hypothesis but precisely in order to improve one's own estimate of such probabilities, Scriven's evaluation of Walter's paper seems an important contribution. Perhaps there are cases left where "coincidence" ~~explanation~~ still "strain the probabilities in a ludicrous way" (Stevenson, 1962, p. 63) but there is no doubt that the strain has been reduced through Walter's finding.

This may suffice to exemplify the undesirable practice of

adding emotional overtones to parapsychological publication.

The term emotion is not entirely suitable and perhaps one should argue instead against "surplus meaning" in publications.

The term "surplus meaning" is borrowed from English & English's (1958, p. 116) discussion of hypothetical constructs and may here be understood as occurring in statements and expressions which are not directly concerned with the events and procedures published and which are not included in standard scientific publications. That is, a footnote or remark expressing appreciation for financial grants or other help received is common practice, but remarks of the kind by Rhine & Pratt (1961, p. 92) that Hansel was invited at their expense but refused to discuss the problems, seems to carry surplus meaning.

There seem to be some statements and procedures where one cannot be certain whether such surplus meaning was included or not, but by pointing to the possibility that it may have been present, such doubtful expressions may be avoided in future.

R. A. McConnell (1947) published an article in the Journal of Parapsychology in which he argued against the claim for the non-physicality of parapsychological events, a claim which was then and probably still is accepted by the editors of the Journal. The editors provided a footnote (1947, p. 111) indicating that McConnell had just completed his Ph. D. In one sense this is entirely factual but in another sense one wonders whether the editors might have left out "just" and substituted perhaps the

year or month if McConnell's article had been more in agreement with their editorial views.

The present writer is not aware of any scientific journals where replies to an article are printed on pages prior to the article to which the replies are made. The Scriven paper (1961)- the published form of an annual Parapsychological Association Banquet Address - was published under the back section of the Journal dealing with parapsychological association matters, and this seems quite reasonable. However at least one ^{previous} banquet address (Murphy, 1958) was published in the main part of the Journal, and consequently one just wonders whether the editors preferred to point to the reply first because they personally may have favoured it more than Scriven's paper.

As indicated above, no reply to this question can be given and even the suggestion of such a possibility may be entirely unjustified. However some more awareness of these possible interpretations may be desirable.

It is not the intention of this writer to present examples of emotional writing by authors opposed to parapsychology but it may be of interest to include at least one example of surplus meaning from a psychologist who has not only shown a critical interest in parapsychology but who has also attempted to deal with such wider issues as scientific orthodoxy and from whom one might expect a more careful analysis (Boring, 1929; 1932).

Yet in his assessment of Soal and Bateman's publication (1954)

Boring (1961, a reprint of a 1935 publication) showed little appreciation of the high order regularities presented in Soal's results. Boring's misrepresentation may be based on insufficient care and study of the subject matter but his remark that Soal's presentation is 'deadly dull' (p. 126) carries certainly surplus meaning in a critical review of a controversial field of enquiry.

The problem of non-physicality

Boring (1929) suggested that psychological movements are at least in their initial stages negatively orientated, i.e. they move against or away from some previously established point of view.

It seems that the postulate for non-physicality of parapsychological events gave a good deal of momentum to the movement of parapsychology and it probably also united a number of diverse interests. The non-physicality movement is directed as the term implies against the acceptance of the physical principle as the only legitimate one in scientific research and (or) against the predominantly monistic view of our time.

What results in parapsychology if not any pre-existing belief, brought about the assumption that parapsychological events are non-physical in nature?

Among the more widely accepted experimental evidence, one can quote the Pearce-Pratt experiment (Rhine & Pratt, 1954) in which the distance between the target and the subject attempting to call the target, was changed without obtaining significant differences in the rate of correct calls. This experiment and similar ones (Rhine, 1935, 1954a) have led to the conclusion that parapsychological events are independent of space and therefore different in principle from physical events as they are understood at present.

Two arguments can be used which make this assumption doubtful. Experimental evidence indicates that a subject's ability to call targets correctly may vary substantially if no changes in the distance between subject and target are introduced (Sosl & Bateman, 1954 ; Broad, 1962). Since a subject with parapsychological abilities may be "aware" in a parapsychological although not necessarily in a conscious sense, of any changes in distance which are introduced during an experiment, it is not impossible that existing space dependent variations may be masked by the subject to an extent that no co-variations can be detected. Schmeidler came to a similar conclusion (1948).

The second argument is, that as the strength of a signal weakens with the increase of the distance of the source of the signal, the ability to detect such signals may, within limits, be independent of this strength.

This is recognised fairly easily if information is transmitted without loss over varying distances. It is likely that on the basis of the above example no claims for non-physicality would be made if all the targets in a parapsychological experiment had been called correctly at varying distances.

However one must recognise a multitude of physical conditions which could create a loss of transmission which does not depend on the strength of the incoming signals, but on such factors as interference (which may be equally disruptive for strong and weak signals) and on any breakdowns at the receiving end such as

interruptions, wrong channelling or wrong decoding processes which could again be quite independent (within limits) of the strength of the incoming signals. This argument is substantially in agreement with Vasilov's views which are reviewed by Ryzl (1963).

The qualification "within limits" should be acceptable to parapsychologists because space independence is claimed only on the basis of the limited range of 'distance-experiments'. On the other hand no limits as to the claimed independence have been found as yet in parapsychology while such limits have been found in physical systems of signal transmissions.

However, if these strong individual variations are taken into account, then it can hardly be said that the borders of the area of space-independent parapsychological events have been tested. That is, although there exists evidence that correct calls (correct in a statistical significance sense) have been made over a considerable distance (Rhine & Humphrey, 1942; Soal & Bateman, 1954, pp. 267-388; Osie & Plonsar, 1956), there is at present no way of settling the question whether the strength of the signal had changed over long distances or not. Indeed suggestive evidence exists which points to a decline in the scoring level with the increasing distance (Osie, 1959, p.290).

Perhaps the claims for non-physicality would not have been so strong if apparent space independence was the only suggestive

Evidence in favour of non-physicality. Perhaps the more direct reason for non-physicality is that at present no generally acceptable theory which explains parapsychological events in a physical manner has been established, although some attempts have been made (Smythies, 1951, Wasserman, 1956). But this in itself is not a sufficient reason and space independence remains doubtful.

Pratt (1961, p. 23) seems to suggest that non-physicality is the distinguishing characteristic of parapsychological events. However, Burt (1961, p.30) indicated that on a basis of dualism, many if not all psychological events are non-physical, yet they would not be called parapsychological. Broad's (1961) reference to basic limiting principles deserves attention as a possible discriminating characteristic but it would always have to be qualified with respect to the "presently known" limiting principles and as indicated above, the experimental evidence does not necessarily justify the conclusion that a conflict with a limiting principle exists.

Pratt (1961) stressed the point that the field should be related to the present knowledge of physicality and not to any future development which one might envisage. Burt's view however, seems to justify that it is hardly possible at present to define clear boundaries of physics because this scientific field is in such a stage of rapid basic development (Born, 1955 ; Heisenberg, 1955 ; Margenau, 1957).

Another argument against Pratt's view would be to list events

which are not understood in present day physics but which are not called parepsychological because of this. Eccles (1953) and Walter's (1953) calculations suggested that the neurological ability to detect and amplify signals is not understood. Although Eccles makes use of parapsychological hypotheses elsewhere (1953, p. 284), the detection and amplification abilities of, for instance, the human brain are seen as mechanisms.

Other evidence for the non-physicality of parapsychological events is given through precognition experiments, certain PK experiments and through the consideration of certain case studies. Precognition seems to be at variance with the established time order. Broad (1962) made it clear that a variety of precognition cases can be discussed.

Using as an example the Tyrrel machine, Broad indicated that in certain cases precognition referred to experimental procedures where the targets to be precognised were only shown after the precognition calls had been recorded, but where definite physical conditions were in existence at the time of the call which could be recognised by the subject (however only by parapsychological result means) and used to infer the targets. Parapsychological results under these circumstances are not necessarily a sign that some future events are correctly perceived but perhaps only a sign that present events are perceived parapsychologically and translated into the categories of targets which will result from the perceived events. Soal's early experiments (Soal & Bateman, 1954) could

be interpreted in this way although in this case there is evidence that the subject depended on an agent perceiving the target shortly after the call was made. The time difference between calls and inspection of targets by agents was nevertheless small and of the order of a few seconds. But even if the time difference is comparatively large there is no absolute necessity to rule out the possibility that correct predictions of targets are based on the correct parapsychological perception of existing events which lead to the future targets. This argument does not hold if a limited amount of indeterminism is accepted and if undetermined systems are used to select the future targets. Random selection devices may be such undetermined systems as Landé's (1958) discussion suggests. There is also a kind of subjective or personal probability which parapsychologists may have in common at least to a degree, which would indicate that as the sequence of causal events which lead to the future targets increases in complexity, for instance, by including complicated mathematical computations (Rhine & Pratt, 1957), the probability that correct calls of future targets are based on present relevant events becomes increasingly small. But again unless indeterminism is accepted there is no point at which one could speak with conviction of truly precognitive events (Roll, 1961).

Rhine & Pratt (1957) introduced precognition as less well established.

On the other hand, time in any physical system is perhaps less understood than space or mass and there are already odd time relationships recognised in physical systems (Jordan, 1933 ; Margenau, 1937 ; Scriven, 1961), which make it less likely that precognition - if it exists - is at variance with present day physics.

Nevertheless the results from precognition experiments should perhaps be interpreted as the strongest possibility that parapsychological events differ from physical events in principle.

The results of PK have been introduced (Hilton, Baer & Rhine, 1943 ; Humphrey & Rhine, 1944, 1945 ; Rhine, 1954a, 1954b) with the suggestion that successful results are independent of the mass involved.

If only quantitative research is considered there is no evidence published that successful results were obtained when stable physical systems were used. A stable physical system may be described as one where over a comparatively long period of time (compared to the time it takes to complete one PK trial) no observable changes occur. An example of such a system which has been suggested for PK attempts (Jerman, 1957)¹⁰ would be a long, transparent cylinder standing in a vertical position on a solid foundation. A small weight is suspended on a thin wire from the centre of the top and a depth micrometer is fixed at the side near the base of the cylinder such that the weight can be touched or almost touched by the steel spindle of the micrometer. The cylinder is otherwise

10. In personal communication with this author.

airtight. By using a simple electric circuit it is further possible to demonstrate and record accurately when the weight and micrometer touch sufficiently to establish an electric contact and when not. If the micrometer is adjusted to a point just before contact is made then one might hope that a subject might wish sufficiently strongly to make the contact by PK. If such experiments had been carried out successfully with different weights then a case for mass independence (within limits) could be claimed although if no clear movements of the weights can be demonstrated, different weights may not necessarily create different conditions for making an electric contact. However at present no one has made claims for successful results under such conditions.

R. A. McConnell (1952), discussing the problem of moving a stationary balanced needle - another example of a stable system - argued that the physical conditions may be unsuitable for PK but also that the psychological conditions are less favourable for PK compared with experiments where dice are thrown, i.e. when an unstable physical system is involved.

At present then, the only kind of experiments which have demonstrated PK (in a statistical probability sense), are those in which unstable physical systems produced changes which seemed to vary with the wishes of subjects but where any small number of such changes could have been due to chance. That is, the kind

of changes which occur when a subject is successful are similar to the kind of changes which occur when the same unstable system produces changes without any attempts to influence it; but if a sequence of changes occurs which seem to be in agreement with the experimental design, i.e. with the subject's wishes and which seems to be in disagreement with the sequence which one would expect from the unstable system alone, then evidence for PK has been established.

In a sense it may be correct to say that under these conditions PK is mass independent because the kind of experiments which have shown any evidence for PK are those where on a basis of purely physical interpretation, no certainty exists that the process of establishing sequences of events (where the mass of the objects in these events is varied) will necessarily lead to different energy requirements.

It is also probably possible to find different physical systems which are mass independent (within limits) as for instance, establishing under certain conditions an electric contact between objects of different mass.

In the more concrete parapsychological example of throwing dice one may expect that any changes due to PK will occur when the dice are in a state of unstable equilibrium or when they are rather close to such states (Nash, 1955). If in a finite number of throws dice ever go through such states of unstable equilibrium

then the energy requirements for changing such states are independent of mass. But this is in agreement with physicality. If actual states of unstable equilibrium are extremely rare, then the energy dependence on mass is still small as long as the actual states of the dice are very close to the state of unstable equilibrium. Nevertheless one would now expect lower results for PK experiments in which dice of larger mass are used. However since it is uncertain how much difference (if any) one should expect, the occurrence of such differences could again be masked through psychological preferences and peculiarities, which are probably here as strong as in ESP experiments, as well as through physical variables which interfere. For example, dice of small mass may have different physical properties (due to differences in material or size) which could create less favourable conditions for an unstable equilibrium position (or a state close to it) that is, for PK.

Other objects which are used in PK experiments such as balls which may run into various channels or coins which are spun, are also likely to go through a state of unstable equilibrium or through states which are close ^{to} it when the initial movement of these objects has slowed down sufficiently.

Forwald (1934a) calculated 'side forces' which he expressed in standard physical units and created the impression that in certain PK experiments so and so much PK success is equal to so and so much physical force. It was pointed out by Nesh (1956) that

the use of physical units is misleading and Forwald (1959, p.123) agreed that his 'side force' is not a real force but a measure which is convenient for comparisons.

The present writer suggested independently in a letter to Rhine that in this Forwald experiment as well as in other PK experiments, no estimation of the physical forces involved can be made (Appendix 1, pp.³⁰⁵⁻³⁰⁸).

Even if the PK experiments cannot be used with any certainty as an example of non-physical mass independence, there is perhaps still the question left whether equal PK success over various distances should be considered more seriously as a case for the non-physicality of parapsychology.

However if states of unstable equilibrium exist in PK experiments any small amount of force will bring about a change and on a physical basis one would not necessarily have to expect different results over varying distances.

Success or failure may depend on the precision of timing and directing energy supplies and not so much on the amount of energy available at various distances.

If as this discussion suggests, PK operates in states of unstable equilibrium or very close to such states, then apart from relevant considerations with respect to the non-physicality postulate, one other consequence of this acceptance must be mentioned.

Forwald (1954b) has argued that certain of his PK results

cannot be explained by air currents because the air currents necessary to produce such changes would have to be quite substantial. The estimation of the necessary air current was however calculated on the basis of shifting a mass from a to b (Fig. 1) rather than on an unstable equilibrium assumption where a shift from the initial position to either a or b (Fig. 2) does not require any force which is anywhere near the magnitude that would be necessary for the first example.

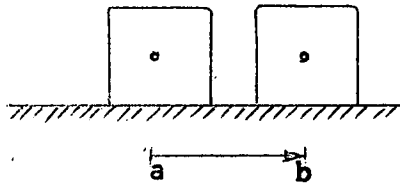


FIG. 1

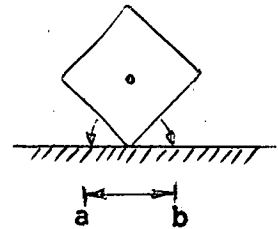


FIG. 2

It seems therefore difficult to reject the air current counter-hypotheses on the basis of the magnitudes of forces involved because they may be extremely small.

It seems still most unlikely that an air current did influence the results because of other difficulties such as directing such a current appropriately and at the right moment.

Empirical data could no doubt be collected by instructing subjects to attempt to shift the dice in to certain target areas by directing an air current in a suitable way. However, in the absence of such information it seems desirable to introduce controls which would prevent any systematic air current influence through subjects or experimenters by placing them into ^{such a} position that where this is impossible, or by covering the apparatus.

Perhaps PK has met with more criticism and with less acceptance than ESP (Wolstenholme & Miller (Eds.) Ciba Foundation symposium, 1956 ; Girden, 1962a, 1962b). Considering such well controlled experiments as those by R. A. McConnell, Snowden and Powell (1955) at Pittsburgh and those by Pratt and Forwald (1958) at Duke, this is perhaps surprising and it may be justified to consider non-experimental evidence which is again linked with the question of non-physicality. Extra-sensory perception of objects or extra-sensory communication between subjects is perhaps traditionally regarded as a mental process outside the realm of physics (Gurney, Myers & Podmore, 1886) and it is interesting to note that physicists and engineers like R. A. McConnell, Forwald and McAulay entered the field by investigating PK. It would appear that on a speculative basis they considered PK to be more of a challenge to physics than ESP.

The present writer would argue that ESP and PK events are equally challenging to physical science because both kinds of events require energies if they are to be explained on a physical basis. The energy requirements for transmitting information (and ESP can be described in this way) have been particularly studied by communication engineers and cyberneticists (Wiener, 1954 ; Parapsychol. Bull., 1960, No. 54), and there is on the basis of present day physical knowledge no evidence which would suggest that information can be transmitted without some energy requirements

and indeed if such an assumption was made it would lead to
 11
 unacceptable results.

It may be appropriate at this point to consider the definition of specific events in parapsychology, such as PK and ESP.

A distinction should be made between PK as a name for a class of possible events (PK changing a stable system p. ⁸¹), and PK as a name for particular events which have been estimated to occur with a high probability (PK changing unstable systems).

If PK is defined as the name for a class of events which are likely to occur with different kinds of probabilities and if some of these events must be considered highly doubtful, then the class name may add some characteristics of the doubtful event to those which are much better established in terms of probability.

It seems that the definition in the glossary of the J. Parapsychol. refers to a class of possible events.

Redefining PK for the kinds of experimental events which have been reported, it may be said that PK is in particular systems a sequential change of events which are not connected by known means

11. Maxwell illustrated the energy requirements for obtaining information by referring to a 'demon' who could set up a perpetual motion by regulating the flow of slow and fast moving molecules without interfering with the laws of thermodynamics if the 'demon' can obtain information about the movements of the molecules without energy expenditure (Wiener, pp. 28-30).

to the subject who desires these changes and particular systems can be described as systems with on-going events of two or more kinds which can be recorded and which are believed to be randomly distributed.

The kind of different events are not changed and usually, because of the limited recording possibilities, such changes if they did occur, would not be registered. What is changed is the sequence of these events.

To establish probabilities for the existence of PK the following quantitative records are compared.

1. The distribution of the events when a subject tried to change the self generated distribution.
2. The self generated distribution of the events which should be empirically established (but a theoretically established distribution may be acceptable). This distribution should be in close agreement with a random distribution.
3. The 'instruction distribution', i.e. the distribution of the kinds of tasks which instruct the subject to attempt to produce corresponding changes in the distribution of events which are recorded.

The situation in ESP is quite similar except that the self generated distribution (2) is produced in the subject. The distribution of targets corresponds to (3) and the change of the

self-generated distribution in the subject to (1).

The randomisation of the targets, i.e. of distribution (3) is a necessary control as otherwise a known connection between (3) and (1) may exist. However, this necessity for randomisation tends to obscure the similarity and the difference between ESP and PK.

Psi has been used to describe the common aspect of both ESP and PK (Thouless & Wiegner, 1946). It may be said that in both cases psi establishes by unknown means a connection between (1) and (3). However, in PK the self generated distribution occurs outside the subject and in ESP it occurs inside the subject so to speak.

To clarify this further, in ESP the subject can be asked to make calls without targets, i.e. the subject is producing a self generated distribution and if targets are then introduced psi will, in successful experiments, change the subject's self generated distribution into one which is similar to the target' distribution.

In both cases the process may be compared with synchronisation, falling in step and resonance.

The above remarks suggest that Jung's synchronicity (1961) may not be irrelevant to parapsychological events. In agreement with Abrams (1958, p. 299) however it seems premature to take the view that strict repeatability is unlikely.

Finally the question of evidence from case material must also

be considered. It was pointed out earlier (pp. 72-74) that on the basis of present information it is extremely difficult if not impossible to evaluate case material quantitatively. However only one or two certain cases of some non-physical form of human survival are sufficient to argue for non-physical aspects of men. But even among the parapsychologists there is no clear agreement on this question. If one is prepared to agree that some of Murphy's (1961), Tyrell's (1961) or Broad's (1962) accounts of cases are sufficiently unique to be accepted as parapsychological, then there still remains the question whether they can be accepted as cases of survival or not.

It seems that the majority of the cases can be explained on the assumption that the 'agent' or medium obtained information parapsychologically which was available at the time of the investigation. The example of a Dutch group spelling out an English poem which was 'in the mind' of an adolescent boy at that time (Murphy, 1961, pp.190-191) seems to provide a good illustration. However the problem is more difficult in the case of cross correspondence (Murphy, 1961; Tyrell, 1961; Broad, 1962). It seems necessary here to assume true retrocognition (H.H. Price, 1960, p.118 ; Tyrell, 1961) or precognition. Either the agent was capable of perceiving information which was only available from some person who had since died, or the image to be experienced was projected towards the agent into the future so to speak.

If precognition is accepted, cases of alleged survival can be explained - perhaps not entirely satisfactorily - on the basis of parapsychological events, without the assumption that any aspect of a human being survived.

Which possibility is more likely shall be left open here but it must be noted that precognition which was previously recognised as being possibly the main argument for non-physicality must be employed in order to exclude survival. If survival is accepted then the non-physicality of some aspects of man becomes more likely.

However it is difficult to estimate how meaningful case studies are in the first place (Walter, 1960 ; Scriven, 1961) and precognition may not necessarily be in strict disagreement with contemporary physicality (Roll, 1961).

How far on the basis of the above discussion claims for the non-physicality of parapsychological events should be continued is difficult to decide. The argument so far was only supposed to demonstrate that there is no strict necessity to claim non-physicality. There may be still enough evidence to suggest that it is quite likely that parapsychological events are non-physical. The estimation of such probabilities is, however, a subjective procedure which will depend on one's estimation of the probability that dualism or monism is correct, on such considerations whether mental events in psychology are fundamentally different from physical events and on what physicality is supposed to mean.

Among the scientists who have indicated their interest in parapsychology through publications, Walter (1953) argued on the basis of neurological processes of the brain and on the basis of physical communications, that parapsychological events must be regarded as non-physical. On the other hand Eccles (1953), equally prominent in neurophysiological research, argued for the existence of a 'mind beyond the brain' but a careful analysis shows that he did not necessarily assume that this mind is non-physical. It seems that this problem cannot be conclusively solved by listing expert opinions.

Accepting some difference (not necessarily a fundamental one) between psychological and parapsychological events, one basic question is whether research is likely to be more productive if carried out in close agreement with psychological methods and procedures or whether indeed such methods hinder rather than help parapsychological progress. One should perhaps restrict the meaning of progress to "progress in research". But this is in actual life no easy decision. Financial help may lead to considerable progress in research and financial help may have been at certain stages, a decisive factor in the continuation of parapsychological work.

Yet the availability of financial support may strongly depend on the political, religious and philosophical issues on which the research could have a bearing. Assuming that some parapsychologist's belief favoured dualism and that financial aid was provided by

donors with a similar outlook, one cannot easily condemn the emphasis on non-physicality, particularly at times when the financial support and hence the continuation of research was by no means certain.

Rhine's earlier claims for non-physicality nevertheless left more room for changes:

"Altogether these findings require either a physical process that is considerably unlike anything that is known today or else a process that is extra physical. But perhaps such a distinction at this stage is not very meaningful" (Rhine, 1940 republished 1961, p. 67).

In 1957 Rhine wrote (p. 247) "by the methods of science parapsychology has conclusively refuted materialism as an adequate philosophy of men."

The most unwavering claim was made only recently in 1962 by Onetto (p.59) : "whatever the outcome may be in this terra incognita, one thing will remain : the non-physicality of psi."

In the light of the above difficulties, the separation of the parapsychology laboratory at Duke from the psychology department in 1935 is at least understandable and was perhaps necessary (Rhine, 1963).

Yet in spite of Rhine's optimistic assessment in the same publication, it seems an unfortunate development that the same laboratory is now in the process of being completely separated from the Duke University, if one reflects on the remarks by the founder of this first university laboratory for psychology,

William McDougall (1937, p. 4):

"But detached research institutions have many intrinsic drawbacks; especially the lack of opportunities for easy intercourse with workers in other fields, a drawback which in the case of an institute devoted to psychical research must be especially serious.

More than any other research worker, the student of psychical research needs to live and work in a community of students and specialists along many lines such as can only be found in the university. He needs contact with them to keep him sane, balanced, critical, yet devoted to the pursuit of problems that may at times seem insoluble, by methods which at times seem trivial, boring, and absurdly inadequate to the immense significance of the problems pursued. He needs to live and work in an atmosphere of many-sided research, critical and skeptical, yet tolerant, understanding, open-minded to all possibilities, one which has achieved some comprehension of the extent of our ignorance and the limitations of our knowledge, one which recognizes that all things are possible save the logically contradictory."

It may not be unrealistic to assume that the financial situation is at present less critical than in the past. Oddly enough this may perhaps be partly due to the materialistically orientated institutes in Russia which provide now a competitive challenge to the Western world.

Perhaps one can return to the question of research in a more optimistic way and assume that financial support is available no matter what postulates are put forward, and then ask again; "is the postulate of non-physicality of any assistance in research?"

It must be admitted that no physical theories have so far been of much direct assistance in parapsychology but non-physical theories have not led to much empirical research either. One argument that seems to favour the assumption of non-physicality is that any attempt of a physical explanation would have to be remarkably complex.

In psychology and elsewhere, the principle of parsimony (Morgan, 1901, p. 53) is useful when in an unambiguous situation a simple theory is favoured instead of a more complicated one.

In particular to devise any unambiguous rules about degrees of simplicity or complexity seems to be a very complex if not impossible task (Schlesinger, 1963). Yet the successful use of the principle of simplicity has perhaps created a climate of scientific expectation which rejects the complicated theory because it is complicated and not because a simpler one of equal explanatory power is available.

In psychology Battig (1962) has argued against the general acceptance of the principle of parsimony.

Gardner (1957) concluded that many pseudo-scientists assume that scientific theories are wrong because they are too complicated or because the mathematics employed are not simple enough.

It is not for this writer to judge whether Dudley should be

classed as a serious scientist or not, but his recent publication (1959) certainly suggests that he objects to physical theories because of their mathematical complexity.

Garan (1963) expressed dissatisfaction with psychological theories because they are too complicated and introduced a simple principle which operates not unlike Lawson's "suction and pressure" (Gardner, 1957, Chapter 6).

There can be little doubt that some physical systems can be very complex indeed. Typical examples are sonar systems of bats and porpoises (Kellog, 1961).

Is the order of complexity substantially different in the ESP or PK situation? Again there can be no clear answer but it does not seem unreasonable to leave the possibility of a physical explanation open. Whether the investigation with serials for very short wave lengths (Turlygin, 1942, reported by Ryzl, 1962, p. 223) was doomed to be inconclusive because transmission by the kind of waves suggested was ruled out through other experiments (Vasiliev reported by Ryzl 1961a, pp. 83-85) shall be left open, but it seems that these kinds of possibilities should be investigated if they are based on sound physical assumptions.

There is perhaps also some danger that the exploration of parapsychological events can be hampered through a too specific physicalistic approach. Konecni (1963) referred to parapsychological events by speaking of "certain phenomena of electromagnetic

communication between living organisms" (p. 66). But this is hardly sufficient reason to postulate non-physicality in parapsychology.

As stated previously this discussion is not supposed to suggest that parapsychological events will be explained in the framework of contemporary physicalism but it is hoped that these considerations have shown that no clear experimental evidence exists which would make the postulate of non-physicality a necessary one.

Numbers as targets

A more specific problem in certain parapsychological tests is the use of numbers as targets.

In a wider sense, this is part of the problem that individual subjects or groups of subjects may form patterns in their calls which are presumably not based on ESP and which are likely to interfere with ESP.

Seriven (1957) suggested that subjects should be tested for pre-existing tendencies to make calls in patterns and those subjects, or in a larger sequence, certain batches of calls should be selected that show the highest agreement with a theoretical random sequence, for evaluation and future ESP tests.

Seriven acknowledged Meehl as the originator of this idea and also mentioned that Roll had started some experiments along similar lines.
12

Presumably those subjects who show strong patterns in their calls in runs without targets, are less likely to change their pattern through ESP, which is usually only a weak influence on a subject's choice of a call.

In the case of number targets this can become more serious since groups of subjects may have certain preferences in common (e.g. the number three). If numbers are selected which form a natural range

12. No relevant publication by Meehl or Roll are known to this author.

(e.g. one to five or one to ten) then it is also likely that groups of subjects may (unintentionally and unconsciously) react to these numbers as if they formed a rating scale and there may be general tendencies to avoid the extreme points on such scales. This is known as the error of central tendency (Guilford, 1954, p.278).

Evidence for this will be presented in the experimental report section (pp. 141-143, 164).

Numbers as targets have become a natural choice since computers have been used for testing (Rhine, 1962a, Smith et al., 1963).

After the Rhine publication the present writer corresponded with Pratt who was then at the Parapsychology Laboratory, Duke University, North Carolina, and pointed out that it would be desirable to obtain information about the choice of cell numbers by the Canadian group because on the basis of a large number of subjects some corrections in expected frequencies could be made for future tests.

It seems desirable also to see whether the numbers generated by the machine could explain the significant negative deviation by non-parapsychological means.

Pratt argued that it is not necessary to show that a particular distribution was random as long as the procedure is clearly described and as long as the procedure is normally expected to produce a random sequence.

Pratt's view can be generally accepted and the probability that a parapsychological hypothesis is countered by a hypothesis based on the distribution of targets is small, and possibly of the same order

as the probability which led to the rejection of the null hypothesis. Nevertheless if it is possible to do so, it would be desirable to exclude even unlikely counter hypotheses.

The argument is fairly simple in an extreme case which may be used for illustration. If a subject who does not know anything about parapsychology is asked to guess numbers from 1 to 5 and if he makes the error of central tendency by avoiding the numbers 1 and 5 to a large extent in a comparatively long sequence of guesses, and if the same subject participates later in an ESP experiment with numbers from 1 to 5 as targets and if the target distribution within this experiment happened to be low for the numbers 1 and 5, then the possibility of coincidence (without any parapsychological meaning) is not ruled out.

Of course it could be argued that precognition operated in the first sequence of guesses or that PK operated in the selection of targets but the above argument is not supposed to demonstrate that no parapsychological events occurred but simply to show that in this case parapsychological events are not necessary to account for the results. Indeed it could be predicted that a subject who shows consistent guessing behaviour of the above kind will guess targets correctly (to a statistically significant degree) whenever a target sequence is presented which strongly agrees with the subject's error of central tendency.

The danger that results are based on such patterns is rather small. Card tests with closed packs exclude the possibility of a target distribution with unequal numbers of targets. In card tests with groups of subjects it is also unlikely that any pre-existing preferences for certain cards will be similar among different subjects.

Any tendency to form calls into patterns and any evidence that some target sequences were comparatively patterned should, in the long run, obscure parapsychological events. The existing evidence for ESP is therefore not likely to decrease if it were possible to analyse the patterns in all previous experiments.

There is some indirect evidence which suggests that particular subjects scored better when their calls were more patterned (Nash & Durkin, 1959 ; Carpenter, 1963, p. 268) but this is difficult to evaluate with respect to the above considerations.

It should be pointed out here that a finite section of a true infinite random number sequence can have any pattern. Hence the evidence that a finite sequence is patterned does not mean that it is not part of a random sequence. But for practical purposes some rule may be convenient which could exclude experiments in which the random sequence showed a comparatively high order of patterns. Such a finite sequence could be tested to see how likely it is that it is part of a random sequence and if the probability that a difference exists (for instance established by chi square) reaches a certain value, the sequence may be rejected.

In a recent ESP test in which computer generated numbers were used as targets (Smith, et al, 1963)¹³ some evidence exists that some subjects responded to the targets as to a rating scale, regarding the numbers 1 and 10 (= 0) as the extreme endpoints of this scale. From the publication only a very small fraction of the total results can be analysed but the 100 calls and targets reported on page 9 show nine calls of 1 and 0 when 20 are expected. A chi square analysis shows that this deviation can be expected by chance less than once in 100 cases (Appendix 2, p. 309).

Whether there is evidence for these kinds of patterns in the total experiment cannot be judged from this small sample. Dagle, a co-author of the experimental report, agreed to analyse the total results accordingly after this writer pointed to the possibility of patterns in the subject's calls but so far no communications about the results of the total analysis have been received.

Evidence for the error of central tendency can probably be found in calls from a test by Kahn and Neisser (1949) which was accompanied by a footnote by the editors of the 'Journal for Parapsychology' (p. 180).

"In a recheck at Duke University it was found that second, third and fourth positions were called by subjects much more often than the first and fifths."

In this particular case the subject's target was the position of

13. In a parapsychological publication Dagle, one of the co-authors, is cited (Parapsychol. Bull., 1963, No. 65), although he is not listed first in the original publication.

a black mark in a sequence of numbers from one to five.

The discussion of problems arising from the use of numbers has probably little relevance to the more usual kind of data in psychology. Guilford (1954) recognised that subjects may introduce errors into rating scales by following particular individual response patterns.

The tendency to form such response patterns can be regarded as weak compared to most psychological stimulus situations. However, if gambling behaviour is studied for instance in connection with the gambler's concept of probability^{J.} (Cohen, 1960), failure to allow for the possibility of individual patterns could lead to misinterpretations of results.

In parapsychology it seems desirable to check whether the results may be distorted through the existence of patterns, at least in tests with open target sequences and particularly when numbers are used. Experiences with evidence from patterns should in the long run, help to extract ESP events more readily.

14. Tests in which the different targets usually occur in unequal numbers.

Summary

The problems discussed here have been discussed critically although this writer supports the general case for parapsychology. Whether this critical attitude is justified or not cannot be easily decided. It could be argued that too much internal criticism might hinder, rather than help, research. It is hoped however that the problems discussed may be of some use in future research.

In statistical interpretation the discussion led to an appreciation of the limitations of conventional significance levels and to some awareness of the possibility of accumulating 'Scheinsignifikanzen'.

At this stage the introduction of any changes in the use of conventional significance levels is not suggested, but at least it has been pointed out how an empirical study could be of help in reaching decisions about these levels. Such an empirical study consisting essentially of a gigantic survey has perhaps little chance of being carried out for parapsychological work alone. However psychologists have taken more interest in related problems recently (Bollas, 1962 ; Rosenthal & Gailo, 1963) and a joint effort may not be entirely out of the question. One difficult aspect unfortunately remains: The survey deals with exceptions and any small scale pretest is not likely to indicate whether a full scale survey will really lead to the desired results.

But even if no survey is ever forthcoming and even if no relevant empirical information on the conventional significance

levels can be obtained, an appreciation of the possibility of 'Scheinsignifikanzen' should in the long run, be of assistance in guiding research.

The pooling problem cannot be regarded as a fallacy however difficult it may be to do something about it on a practical level. It seems that this problem should at least be considered when similar experiments are carried out which could be regarded as repetitions.

Schneidler's demand (1959) to vary experiments (which seems justified) and the general desire to repeat experiments for the sake of confirmation, can be combined since the amount of permissible change in a repetition can be arbitrarily defined. However the problem of predicting the outcome of experiments and the problem of pooling should be considered seriously.

The problem of repeatability has been considered by many experimenters over a considerable period. It may be unrealistic to expect a major advance from concentrating on more or less disguised experiments, but it seems to be one approach which ought to be tried out.

The discussion of emotional overtones in journal publications suggested that surplus information is undesirable and a hindrance in the evaluation of the problems published.

Experimental evidence does not lead to a clear conclusion that parapsychological events are non-physical. Any strong emphasis on the possibility of non-physicality seems unjustified, particularly

since parapsychological events which are fuzzy and indistinct themselves do not stand out against a precise and clear background of physicality.

Rather are they submerged in a pool of rapidly shifting physical concepts.

Numbers as targets can be used but under various test conditions target numbers may lead to specific patterns which could obscure parapsychological results. In a particular experiment patterns may suggest a counter hypothesis but recognition of pre-existing tendencies to form patterns should, in the long run, extract rather than submerge parapsychological evidence.

Part II The Experimental Report

Introduction

The present writer had attempted previously to list and evaluate the possibilities that might advance research in PK and to some extent in parapsychology generally. These earlier considerations were part of examination requirements, and not published, and those possibilities which are of special interest¹⁵ for the experiments reported here will be summarised.

Standardised test procedures are accepted as desirable in parapsychology (Rhine & Pratt, 1957). However even if test situations are as standardized as in some psychological tests (Wechsler, 1958) considerable differences in test results occur in the parepsychological situation. Attitudes towards parapsychology may also have a bearing on the occurrence of parapsychological events (Schmeidler & McConnell, 1958). Successful subjects are usually not aware when they are successful and when not (Rhine, 1958). It has so far not been systematically tested but there is at least no evidence which suggests that parapsychological events could not occur when subjects do not intend to obtain such results.

15. This summary is based on an earlier paper but is not just an abstraction of it. However to ensure that Rule 9 (Rules Ph.D. p. 67, Univ. of Tasmania Calendar, 1963) has been fulfilled, a copy of the relevant section of the original paper will be included (Appendix 3, pp. 310-330).

There is some suggestive evidence that parapsychological events may be present in suitable life situations. Different results in the original Millikan oil drop experiments (Dampier-Whetham, 1929, pp. 387-388) may have been due to different expectations of the experimenters and due to their unintentional and unconscious

16

PK (McAulay, 1956).

From these considerations it follows that one line of possibility that could lead to more reliable results, is to present a parapsychological test in a disguised form, as a psychological test and to remove the experimenter in order to eliminate the experimenter subject variable and to replace him by some automatic set-up.

However it must also be recognised that a suitable experimenter may be a necessary condition for a successful test. (This was discussed in some detail on p.6-7). It may at least be reasonable to suggest tentatively that some motivational aspects of the experimental situation can be regulated by providing suitable reinforcements automatically at each trial with little delay in time. One previous attempt of this kind (McElroy & Brown, 1950) was at least not discouraging.

Recent discussions (Scriven, 1961; Woodruff, 1961; Werthmann, 1961) suggested that key experiments should be tried again, or

16. In personal communication with this author.

in other words, that it is important to demonstrate parapsychological events in experiments which give statistical significance of a very high order. While it is desirable for theoretical considerations, to obtain highly significant results over a comparatively small number of measures or comparatively fast, (see p. 41) experiments which give results which are just significant but which can be repeated by other scientists, are equally desirable, as long as there is no high scoring test which can be easily repeated.

By exploring the above possibilities it is hoped that some small contribution towards this end may be made.

The kinds of experiments which were explored were: GESP experiments which were disguised as subliminal perception experiments, GESP experiments with targets (favourite pieces of music) which might have a strong emotional and motivational relevance to the subject, automatic PK experiments which were disguised (as psychological experiments with physiological measurements) and undisguised, and where reinforcements were introduced shortly after a physical event which may have been influenced by PK.

The design of the PK apparatus included a number of features which makes it possible to consider further possibilities (Appendix 3, pp. 310-330) but which as yet have not been properly tested (p. 236).

The question whether parapsychological processes operate in daily life without the awareness of the participating subjects is

of some interest and was considered prior to the design of some of the experiments reported here.

It seems desirable to go beyond the analysis of single cases which do not clearly fit the category "without awareness".

The most remarkable efforts which throw some light on this question were made by Cox (1957) who investigated from available records the sex of the 5th child in families where the preceding four children were of equal sex.

Cox's assumption was that in the majority of cases parents would desire the opposite sex and that by parapsychological processes the opposite sex may in fact have occurred more often than would be expected by chance.

Unfortunately it is difficult to decide what distribution of the sexes is to be expected for the 5th child as was pointed out by McWhirter (1957). There was also a small amount of ambiguity in the available records. Nevertheless Cox's results are suggestive enough to assume that the background probabilities ($p = \frac{1}{2}$) for disguised tests are favourable.

In another study Cox (1956) investigated the records of railway companies to see whether any difference could be found between the number of passengers travelling on trains which met with an accident and the number of passengers travelling on similar trains on comparable days. Obviously there are a number of possible errors which make it difficult to assess this study.

This was recognised by Cox who plans to use the number of cancellations on commercial airlines in a future study of "unconsciously applied ESP" in connection with accidents. Since airline records are more precise and since airline disasters are survived by few passengers, some of the difficulties of the published report may be overcome.

Cox's work deserves attention also because his investigations proceeded into areas which were promising on various grounds.

Whether a human's relationship to life and death is accompanied by parapsychological processes can still not be answered but it is this sort of question that should be investigated and Cox has shown that such investigations are possible in principle.

The present writer had suggested to Rhine that a particular kind of lottery may be investigated for parapsychological processes without awareness. Although prize money is less a direct matter of life or death it is not too far out of those categories of events which Cox investigated in an exploratory way. The particular lottery in question has several advantages from a control point of view and the principal approach will be outlined here.

The lottery which seems most suitable is the German Zahlen Toto which is handled very much like various football or soccer lotteries except that here all references to actual sports events

are omitted and participants are simply asked to select six numbers between one and 60. The prize money is distributed according to some formula such that one correct entry of all six numbers receives a high prize which will be reduced if several correct entries are received. Those who have five numbers correct obtain a smaller proportion etc.

Since the targets are numbers, some difficulties which were discussed earlier (pp. 102-107) will have to be overcome. However this lottery is popular and played once a week by a considerable percentage of the total population. Hence it should be possible to obtain information about the distribution of numbers which were entered by the betting population and which did not win, and to compare them with the distribution of numbers which were targets so to speak. Not all numbers are likely to get the same expected frequency (on the basis of previous entries) but once such expectancies are established it should not be too difficult to analyse, for instance by chi square, whether the total betting population chose six particular numbers significantly more often when they were the 'targets' than when they were non targets in previous games.

At present it is difficult to assess the probability that

17. The distribution of prize money may have changed in detail however the principal procedures are likely to remain unchanged.

parapsychological processes are active in situations where participants are unaware of such possibilities. However, the discussions so far seem to encourage experiments which make use of these possibilities which in turn make the results of certain parapsychological tests more reliable.

Apart from some previous attempts in this direction by this writer which were encouraging, there seems to be no published evidence of disguised tests. However some experiments were successful in which parts of targets were unknown to the subjects in a combined ESP, PK test (Osie, 1953). However in this case subjects were still aware of the parapsychological nature of the experiment and they knew that they were supposed to wish for one of six numbers from a die face even if they did not know by sensory means which particular number was desired at particular trials.

In 1962 Tort gave a report at the Parapsychology Laboratory, Duke University, North Carolina, USA about an experiment which seems to have more pertinent similarities to the research carried out by the present writer.

18. As far as is known Tort's results have not been published except for a very short note which probably referred to the above experiments in the J. Parapsychol., 1962, p.265.

Tart used engaged couples as pairs of subjects in disguised experiments. One (subject A) was placed in a soundproof room and physiological measurements of blood volume, EEG and GSR were recorded continuously. The other (subject B) was placed in another room with electrodes connected to him. B received harmless but presumably unpleasant electric shocks in a random time sequence. Physiological recordings were also taken of B. A was not presented with any sensory stimuli. The whole experiment was disguised to both as a physiological experiment. Tart found some slight changes in the physiological measurements of A at the times when B received a shock. Chi square analysis showed a significant difference between some kinds of the physiological recordings of A at the shock times of B and the physiological recordings between shocks.

There are some possibilities that these results do not necessarily support a parapsychological hypothesis. The significance levels were not extremely high and of the order of $p=0.01$. It is not quite clear whether any particular changes were singled out for testing the ESP hypothesis and a significance at $p=0.01$ may have to be further reduced if several possible independent ESP hypotheses were tested simultaneously (pp. 47-48).

The selection of a time period for analysis in relationship to the time of shock as well as the selection of particular changes in the physiological recordings is a somewhat arbitrary procedure.

Appropriate controls could have been tightened up in a second experiment. For technical reasons such as availability of apparatus, Tart could not continue with his work.

When highly sensitive recorders are used, the possibility of interference through non parapsychological means when the electric shock is given is not completely out of the question. However it seems that Tart was careful in his apparatus set up. The isolation of subject A and of the recording equipment from any nonparapsychological interference is of course a necessary prerequisite to make this kind of research meaningful.

Tart's experiment involved parapsychological possibilities which were directly disguised. ESP tests with children as subjects however, are also disguised, at least in the sense that younger children are not aware of the controversial nature of the tests.

As has been discussed elsewhere (pp. 59-61), these experiments with children were comparatively successful.

From this short discussion it can be concluded that on the basis of other research, the background probabilities for disguised tests are at least not unfavourable.

The open presentation of a parapsychological test may have motivational advantages with certain subjects. However it ought to be possible to get the right sort of motivation by other means or at least this should be tried.

It is possible however that in undisguised tests subjects

have more initial information which might put them in a more favourable position.

Without precise knowledge as to what actually happens in ESP or PK the information that target No. 1 is in room X in building Y at time t_1 , does not seem to be of much help. Similarly in the case of dice it is difficult to perceive their movements clearly, particularly if more than one is involved - although Roll (1961) argued that not necessarily more than one die was ever influenced at a time - and it is more difficult if not impossible to judge at what time they should fall over towards what side.

Any disadvantages in this respect in the disguised test can be further reduced by supplying similar information without disclosing the real nature of the experiment.

In the disguised PK test for example, (p. 208) subjects knew that the ball entering the left hand channel would produce a shock and that the ball entering the right hand channel would produce a slide instead of a shock. They also assumed that they could influence the distribution to some extent through changes in their skin resistance which depended on their emotional changes from trial to trial.

The plan to investigate a number of research possibilities was only drawn up on a broad scale initially. Further details were formulated on the basis of the results from some completed

sections and on the basis of special opportunities or the lack of facilities and on other research findings.

The experiments will be described in detail but a summary of the various investigations may be of some help to show the connections. In 1960/61 University of Tasmania:

1. Adaptation of an adding machine as automatic recorder.
2. Construction of a ball 'P.K.' machine.
3. Testing of apparatus including tests of randomisation.
4. Testing of individual subjects in disguised P.K. tests.
(Only a limited number of the possibilities of the machine was used but responses were reinforced through stimuli.)

Parallel with the above research other possibilities of disguised testing were explored during the same period:

1. Setting up of available apparatus for 'subception' tests.
2. Developing procedures under which disguised GESP tests could be tried out.
3. Group GESP tests disguised as 'subception' tests.
4. Individual GESP tests disguised as 'subception' tests in which wrong responses were 'punished' by mild electric shocks.
5. As under 4. but with three different experimenters who themselves were not aware of the parapsychological nature of the experiments.

In 1961/62 (Parapsychology Laboratory, Duke University, N.C., USA):

Disguised tests were of uncertain value in this setting since it could not be assumed that subjects do not become aware of the parapsychological nature of the experiments.

Although the main part of the PK apparatus was brought to Duke some difficulties such as finding a suitable adding machine, transformer and base for the apparatus, delayed experimentation until 1962.

The situation was more difficult with respect to subception experiments. No automatic timers or slide projectors were available and an attempt to build a mechanical timer with practically no tools or workshop facilities did not lead to a usable set up.

Opportunities existed to carry out physiological measurements and the author attended classes in EEG interpretation at the Duke hospital and assisted Dr. Tenny at the Veterans' Hospital in Durham and at the parapsychology laboratory, in carrying out exploratory ESP tests while physiological changes (EEG, GSR and blood volume) were measured.

The recording equipment for the physiological measurements at the laboratory did not work satisfactorily at times and this work can only be seen as a training for possible future work.

However one experiment was conducted in co-operation with Wadiah Saleh who was at that time a member of the parapsychology laboratory.

PK dice tests with the motor driven rotating cage were carried out and served as an introduction to participating subjects.

Finally a GESP music test was carried out as an attempt to create more stable test conditions and during the same period the PK machine was put back into operating order. Subjects were tested with it in an undisguised setting but without reinforcing stimuli due to the lack of an automatic slide projector and suitable electric shock facilities.

A GESP group experiment disguised as a subliminal perception test.

The initial work was carried out before Nash's publication on relevant experiments (1961, p.278) was available.

A comparison of actual subliminal perception results with parapsychological results is obviously desirable but such attempts were given up in the initial stages because the apparatus available was not sufficiently accurate for subliminal research.

The only interval timers which were available during these tests in 1961 had as their shortest time interval 0.01 seconds. This interval was not considered reliable but no data concerning the errors were available.

Since most subjects can see a considerable amount in 0.01 second, stimuli had to be presented more or less out of focus in order to go below recognition thresholds.

By turning stimuli out of focus it is likely that variables are introduced which may interfere with the general notion of subliminal perception.

But it seems desirable to restrict this term to the perception of stimuli which are thought the reduction of projection times brought below some point of awareness (which may be defined arbitrarily) and not through reduction in clarity or illumination.

To overcome this problem the present writer designed an apparatus for very short projection times (Keil, 1962) which allows simple stimuli to be projected under optimal viewing conditions but for sufficiently short time intervals to be below the recognition threshold (Appendix 4, p. ³¹31). The projection times can be varied by very small amounts in a controlled way. The stimulus is projected with projector No. 1 (see Fig. 3) through two holes cut into two disks which rotate in opposite directions with a slight difference in their speeds. As a result of the speed difference the projection of the stimulus occurs only (depending on the gears used) every twenty-eight revolutions, hence sufficient time is allowed to turn projector No. 1 on and off.

Projector No. 2 provides illumination between stimulus presentations but the illumination is cut out (by the use of fins) precisely at the moment when the actual stimulus is presented. The fins however cause a slight flicker between trials.

A prototype of this machine was built in 1961 for an industrial firm and the apparatus was not completed in time for subliminal or disguised ESP (or both) research. At present a

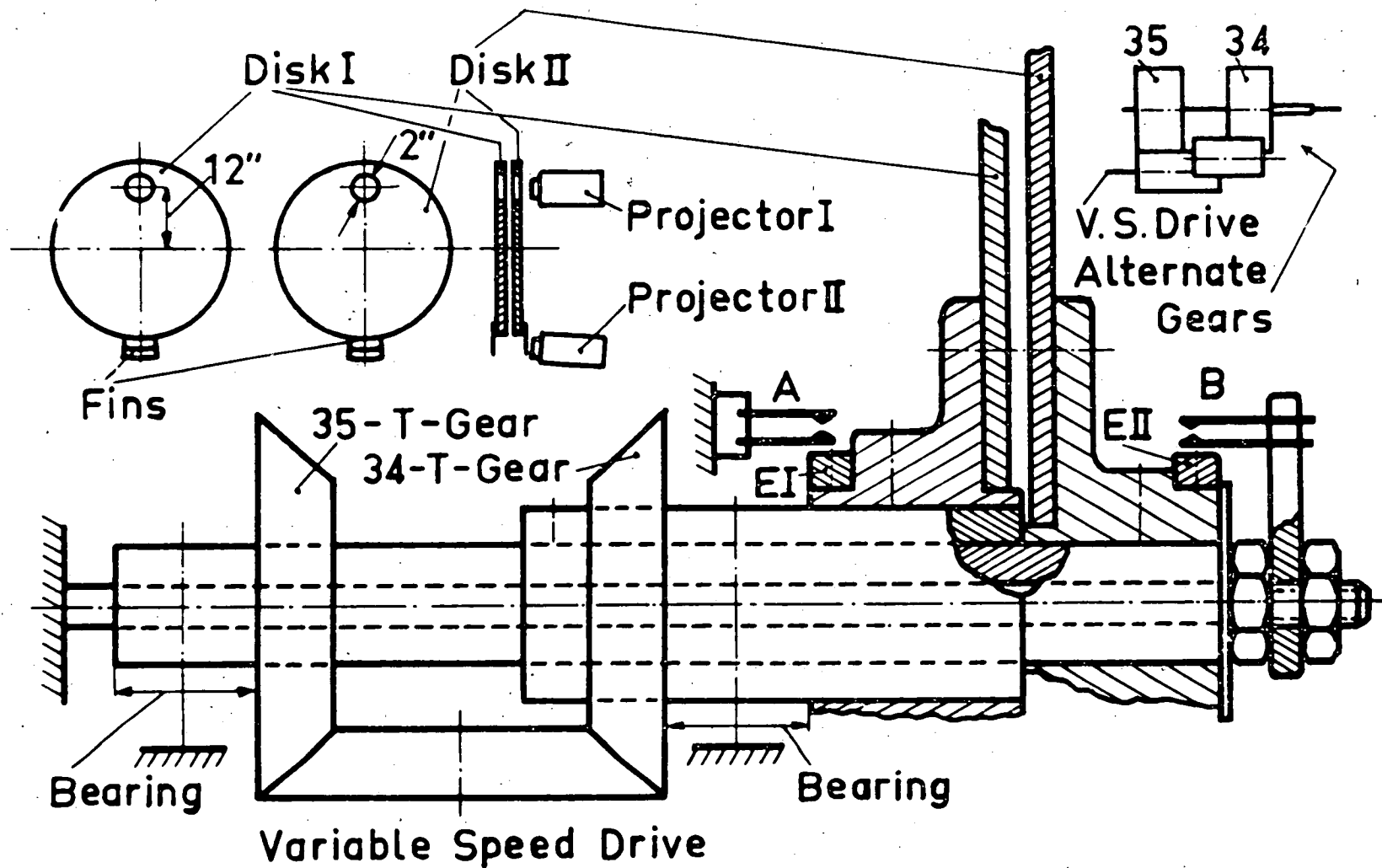


Fig. 3. An apparatus for very short projection times

second apparatus of this kind is being built for the Psychology Department, University of Tasmania, and it is hoped that it can be improved such that the flicker can be eliminated without reducing other functional possibilities of the apparatus.

In the GESP experiment one interval timer was used to flash a stimulus for the 0.01 second setting on a screen. Numbers were used as stimuli in the form of three letter words (ONE TWO THREE FOR FIVE) and when the subjects were asked to guess numbers on the GESP part of the experiment the three letters GAP were presented throughout while the subjects expected that any one of the five numbers was presented.

The purpose of this experiment was mainly to see whether under these test conditions any GESP results can be obtained. The GESP hypothesis was to be tested by analysing the direct hits only. Other features of interest in the results were to be investigated and analysed but this was not to be taken into account when the GESP hypothesis was to be tested.

The experimenter expected a positive deviation but a one-tail test was not specified in advance.

Before a final procedure was adopted, various aspects of this test were tried out without considering possible ESP results.

In this setting as well as in others (i.e. using a PK apparatus) a comparatively large number of test runs was necessary before technical conditions were sufficiently suitable to attempt any experiments.

Usually this is accepted as methodologically sound provided it is clearly laid down when the experiments start and stop. However on account of the perhaps more extreme considerations (pp. 50-51) it might be desirable to pause here and to see whether this is methodologically different from example two (p. 51) or whether any experimental results should be interpreted on the basis of the total number of trials (no matter whether in the experimental situation or not). The latter case however does not apply because the trials of the experiments were prior to the experiment announced as "experimental trials". This is equivalent to the discussed case (pp. 53-54) in connection with example one (p. 49) where out of 10 gamblers, one particular candidate is preselected to win. In this case his chance to win is independent of the number of other gamblers attempting to win.

Or to return to the experimental situation in parapsychology: if any possibility exists that some of the non-experimental trials could be used to make claims for parapsychology, then the experimental trials must be evaluated on the basis of the total number of trials. However if, as was the case, the non-experimental trials are clearly distinguished as being of no potential evidence (as far as testing the GESP hypothesis is concerned) then the experimental trials can be considered independently of the non-experimental trials.

In the initial non-experimental runs it was found that a stimulus which was sufficiently blurred to be unrecognisable to

a few single individuals who knew the stimulus in advance, became recognisable to some subjects in a group when they viewed the stimulus repeatedly without knowing or expecting it.

This could be of some interest to perception in psychology but most likely the recognition was in this case not based on repetitions of a stimulus which was practically the same throughout but on variations in the timing of the stimulus. That is in every 20 contacts which the timer made about one may have tended to "stick" slightly longer thus producing an increased stimulus time.

Since the stimulus time could not be reduced it was necessary to put the stimulus further out of focus so that the stimulus could not be recognised even if projected for some length of time.

But by now some dissatisfaction with the tests among the subjects became apparent through responding with written words like 'light flash only' when they were supposed to select a number from one to five. This was based on the repeated viewing of a light flash which at times was sufficiently long to expect the stimulus letters to stand out clearly.

Most subjects still followed the instructions and wrote down numbers but it seemed that more than just emphasising the relevant instructions was needed.

It was found that a short sequence of 15 'subliminal' trials

could be used to break down resistance in a group. It was also found that by limiting the subject's response time to five seconds, subjects had 'no time' to feel too concerned about not seeing the stimulus. Finally the following procedure was adopted.

Instruction sheets were handed out to the subjects which read as follows:

Research on Subliminal Perception

This is not a demonstration but an experiment where results are supposed to add new knowledge to this field.

You are sincerely asked to co-operate to the fullest extent.

Introduction

You may have heard of subliminal advertising, e.g. during a motion picture show, advertisements are projected on the screen, but only for such short time intervals that nobody could recall having seen them. However the sale of the advertised product increased significantly.

More reliable results were obtained when in conditioning experiments the stimulus was first presented for time intervals sufficiently long to ^{be} perceived by all subjects. If the stimulus was shown coupled with an electric shock, subjects showed later the conditioned response to the subliminal presentation of the stimulus alone, i.e. when

they were not aware of recognising the conditioned stimulus in a series of other stimuli also shown for short time intervals only. To detect the conditioned response the change of skin resistance was recorded.

Purpose of Experiments: Some evidence exists which indicates that subliminal perception depends on the intelligence of subjects, e.g. a group of university students should obtain better results (recognize more stimuli) than a group of comparable adults with lower intelligence.

Demonstration Experiment

To avoid any recording errors the stimulus words are numbers. To obtain equal numbers of letters some letters have been omitted.

Instructions:

- ONE - No. 1
- TWO - No. 2
- REE - No. 3
- FOR - No. 4
- IVE - No. 5

will be presented at random.

Use the booklet marked Pre-test only. After each presentation the light will come on for 5 seconds. Write the number of the stimulus which is now in the projector in your booklet. Write down all stimuli. Do not leave out ANY. If you have no idea which one you have seen, write down the one which comes into your mind most readily.

Call it guessing if you like but it seems that the correct response will be given even if the stimuli are not consciously recognised. Do not attempt to work out any order or pattern. This is impossible anyway but such attempts may interfere with the experiment. Fold over the page just used in your booklet and get ready for your next trial. The light will be switched off shortly before the next stimulus is presented. There will be 15 trials. The trial number will be announced before each presentation of a stimulus.

Controls: It is most important that there is no communication of any kind between subjects of this group during the experiment. Please do not smoke. This may impair the view of some subjects.

Experiment

The projection times will be shorter and it will be much more difficult to recognise the stimulus word. There will be 100 trials. You have four booklets of 25 pages each to record your responses. Your four booklets are numbered through from 1 to 100.

For technical reasons there will be a short interval after each set of 15 trials. Instruction and control as for Demonstration Experiment.

Prior to the demonstration experiment the five stimuli were projected for 0.5 seconds each.

In the demonstration experiment the five stimuli were presented at random for the 0.01 second setting of the timer and somewhat out of focus. The out of focus setting was such that about one in 20 to 30 subjects was able to recognise all stimuli correctly. Under these conditions it was also usually found that some subjects had only chance results or less.

At the end of the demonstration test the correct sequence was shown by opening one half of a 60x1" folding blackboard. The sequence was written up there. Subjects were asked to score their own results. The extreme results which were found to be present in all groups tested so far were used to 'demonstrate' that at the speed used for the demonstration experiment some could still score all stimuli correctly while a few were unable to score above chance. It was then emphasized that "in the main experiment the projection time will be considerably shorter and consequently it will be much less likely that any of you will be aware of the stimuli presented, nevertheless you must try your best to respond with that stimulus which comes most readily to your mind."

The presentation of stimuli and the control of the time sequence of the events making up one trial was automatically triggered off by five interval timers which performed the following actions.

Timer No. 1, 2 seconds, Action 1. Room light off. This is signal that stimulus will be presented soon.

Timer No. 2, 0.01 seconds, Action 2. The neutral stimulus (N) is projected on screen.

Timer No. 3, 0.5 seconds, Action 3. Slide change mechanism of projector is activated. Stimulus No. 1 is moved into the projection position but not projected.

Timer No. 4, 5 seconds, Action 4. Room light on. This is signal to record trial No. 1. At the beginning of this period trial No. 1 is announced.

Timer No. 5, 0.5 seconds, Action 5. Room light off. Slide change mechanism is activated. Stimulus No. 1 is returned into the slide magazine and (N) is brought into the projection position.

Each slide magazine can take 30 slides. 15 slides of the 5 stimuli were arranged in random order (according to a random number sequence) occupying the second, fourth, sixth, etc. positions in the magazine. The neutral stimuli (N) consisting of three letters GAP, were placed in the first, third, fifth, etc. positions, i.e. 15 neutral stimuli were placed in each magazine. The slides were obtained by typing the required capital letters on semi transparent tracing paper. This paper was cut such that the letters appeared approximately in the centre and the papers were placed into 2 x 2" glass frames which are commercially available.

The interval timers can switch one circuit on and one off during the test time intervals. At the end of each interval each timer can be wired to start the next timer and so on to the fifth timer which starts the first timer again.

In the above sequence timer no. 1 does not actually switch anything else during the set period. It represents a 2 seconds pause before it starts timer no. 2. Both timer no. 3 and timer no. 5 activate the automatic change mechanism of the projector. Thus during each trial slides are changed twice. Hence while subjects record their responses (action 4) the target stimulus is in a special position (the projection position i.e. removed from the magazine) but never actually projected.

In the demonstration experiment the sequence of action was the same but the random sequence of 15 stimuli was represented by two slides per stimulus. That is the sequence REE, ONE, FOR was represented in the magazine by REE, REE, ONE, ONE, FOR, FOR, but only REE, ONE, FOR, was projected.

Subjects were therefore trained in the demonstration experiment to recognise the two seconds darkness as a signal for the stimulus to be presented at the end of this period and to get used to recording their responses within 5 seconds.

The automatic sequence of the five timers could be interrupted through a switch between timer no. 5 and timer no. 1. This was done at the end of each magazine of slides. After the new magazine

had been placed into the projector the first N stimulus was moved by hand operated mechanism into the projection position and after closing the interruption switch the new sequence of 15 trials could be commenced by starting timer no. 1.

The announcement of the trial numbers was made by the experimenter or a demonstrator (senior student employed by the psychology department), in a monotonous voice. Announcing the trial number helped to avoid recording mistakes.

During the actual experiment the trial numbers were announced by a demonstrator who did not know the target sequence.

The experiment was carried out in the psychology laboratory and lecture room in the basement of the former domain university of Tasmania, Australia (Fig. 4).

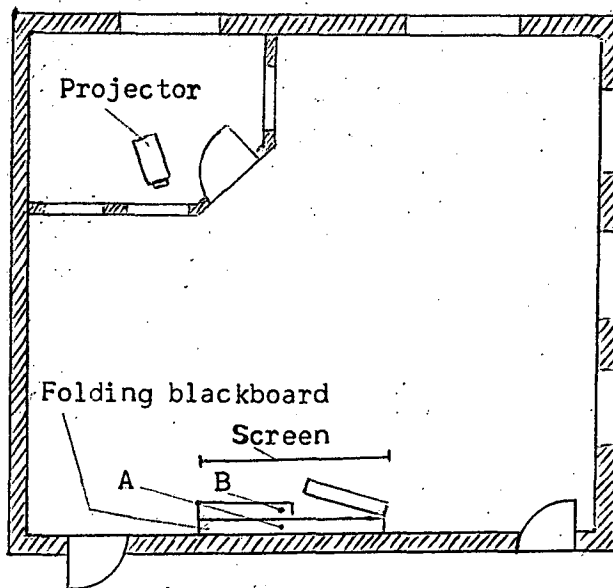


FIG. 4

Ground plan of test room for the GESP group experiment.

Not to scale.

The room accommodated approximately 30 to 40 students.

One back corner of this room was completely boxed and glassed off and to some extent sound proof. Windows were double windows fixed permanently in the walls. This boxed off room could only be entered from the lecture room and it was used during the experiment to house the projector and timing equipment. A portable projection screen was placed in front of the blackboard (see Fig. 4).

Under these conditions subjects had no possibility of obtaining information about the targets by non-parapsychological means. Slides in the magazine or in the projection position could not be identified by sensory means.

Even if subjects had attempted to look through the double windows and straight through the projection lense they could only have recognised a target when the projector lamp was on. This was never the case for actual targets. There was also no evidence that subjects attempted to obtain information by these means.

The lecture room could be almost completely darkened by appropriate blinds, and during the evening (when the experiment took place) it was practically blacked out.

Four portable lamps were positioned high above the subjects to give sufficient illumination to all subjects to record their responses. The lamps were connected to the timers and were switched on and off as indicated (p. 133).

Thus an effective control was provided for the response times which prevented without much instruction, mutual communication between subjects because there "was not enough time for it".

A group of 40 Psychology I students was tested in two successive sessions with approximately 20 students each. The target sequence was the same for both groups. The correct sequence was not revealed after the test to either of the two groups. The Greville correction (Pratt, 1954) was employed for the combined group of 40 students.

The 40 students represented about two thirds of the total number of Psychology I students and included all those (except students who were absent for other reasons) who had not participated in previous non-experimental runs.

In addition to the target sequence in the projector the correct target sequence was also written up on the right hand blackboard and then covered by folding the blackboard over. No reference of any kind was given to this blackboard sequence but it was to some extent implied by using the left hand blackboard in a similar way to write out the correct sequence of the demonstration experiment stimuli and by revealing the sequence at the end of the demonstration experiment.

To avoid any possibility that subjects might open the right hand blackboard without the experimenter's knowledge, a small

thread was tied underneath the boards between two pins approximately at A and B (see Fig. 4). This 'seal' which was not noticeable without close inspection and all other preliminary arrangements, were completed before any subjects entered the lecture room.

The portable projection screen in front of the blackboard also prevented the board from being opened accidentally.

During the two successive experimental sessions the experimenter was present throughout and the 'seal' was found to be intact at the completion of the total experiment.

After the subjects from the first session had completed their response booklets they were asked to leave quietly and not to discuss any aspects with incoming students. However if such discussions took place they should have no bearing on the analysis (including the Creville correction) of the total group.

Both groups were told after a few days that their results were most promising and that it would be most interesting to compare them with slightly different individual tests. Results were not revealed until after the completion of all individual tests (p. 150).

Results of group test.

The 40 subjects completed all their responses except for two single responses which were left blank by two subjects. Since on the basis of non-experimental tests no omissions were expected in the final experiment, no provisions for excluding

any subjects were made prior to the experiment and the recorded responses of these two subjects were included for analysis (the two omissions were, of course, scored as non hits).

Since only two omissions occurred in 4000 responses, results were analysed on the basis of the full number of responses. For a positive deviation this amounts to a very small error which is unfavourable to the GESP hypothesis.

A positive deviation of 80 (direct hits) occurred in the equivalent of 160 runs. If such a deviation had occurred in independent runs the critical ratio would be 3.16 or the probability that such a deviation could occur by chance would be 0.0016.

However since subjects called the same target the Greville correction had to be applied (Prett, 1954) and the critical ratio was found to be 2.38 (Appendix 5, pp. 334-337). The probability that a deviation of 80 occurred by chance is 0.018 or less than 0.02.

If the 0.01 level is used to separate significant from insignificant results the null hypothesis must be accepted and no claims for GESP can be made on the basis of these results.

If the 0.02 or 0.05 level is used instead, the results can be called marginally or just significant.

The deviations (direct hits for the total group) were distributed in the four quarters as follows (see Fig. 5):

First, +45; second, +8; third, -3; fourth, +30.

(Each quarter consists of one run or 25 targets per subject).

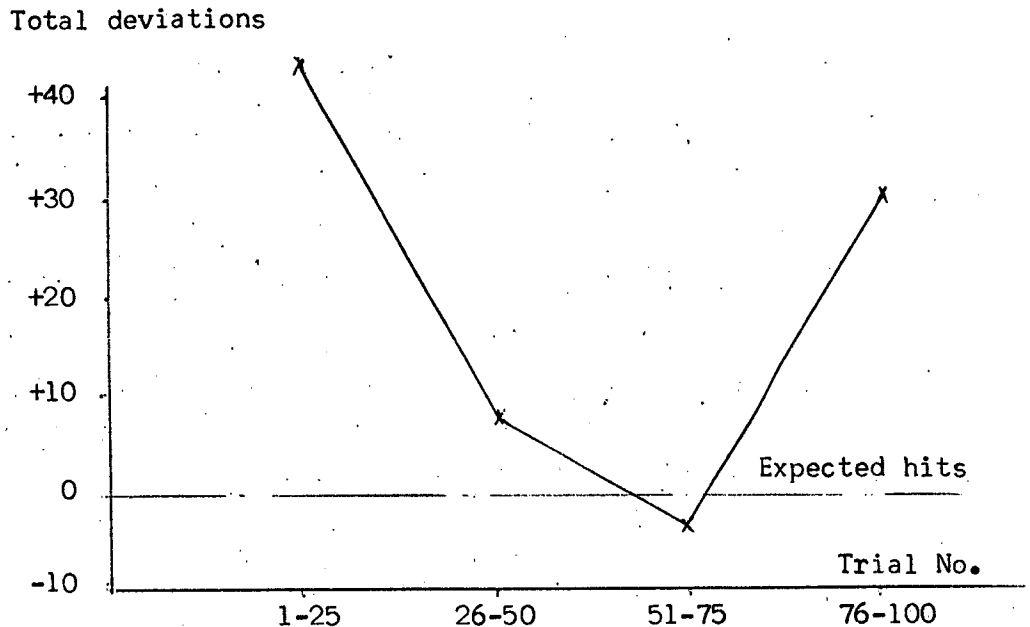


FIG. 3. Quarter distribution of direct hits in group test with 40 subjects. 160 runs, multiple calls, total deviation = +60.

When the Greville correction was applied it was found that the difference between the 1st and 3rd quarter could have occurred by chance with a probability of 0.055 and the difference between combined 1st and 4th and combined 2nd and 3rd with a probability of slightly less than 0.05 (Appendix 5, pp. 334-338).

Discussions:

Whether the results can be formally accepted as significant or not shall be left open here. The author had no fixed dividing line prior to the experiment except the 0.05 and 0.01 level

from general experimental psychology. That is, most results may be considered seriously if they are based on $p \leq 0.05$ but the possibility of a No. 1 error must be kept in mind particularly if $p > 0.02$.

Since the critical ratio dropped considerably once the Greville correction was applied, some evidence exists that the subjects responded in a similar way. From subjective observations it seems to this experimenter most unlikely that subjects communicated with each other.

From the distribution of calls with respect to the five target numbers, it becomes apparent that the subjects responded to the five targets as if they formed a rating scale. That is they avoided the extreme 'ends' one and five (see Table I) and displayed the error of central tendency (Guilford, 1954).

TABLE I

Total number of calls per target figure

Target figures:	1	2	3	4	5
	694	915	904	875	610

Expected number of calls per target = 800

If the 4000 scores had been made independently (i.e. by 4000 subjects) it could be immediately concluded with the aid of chi square that the actual distributions differed significantly from the expected distribution. (For target 5 alone the difference

between f_o and f_e squared $\approx 36,000$ which divided by 800 is ≈ 45 but for four degrees of freedom $p = 0.001$ for chi square = 18.5 ; i.e. $p < 0.001$). Since 100 calls were made by each subject it is possible that this difference could be due to a few extreme subjects. Since the error of central tendency is of some interest in connection with other discussions, it was also tested whether there was any significant difference between the number of subjects whose combined end scores (for target 1 and 5) were below the expected value, and the number of subjects whose combined end scores were equal to or above the expected value.

By chance the number of subjects in the two groups should be equal or slightly higher for the latter group if any subjects have numbers of calls which are equal to the expected figure.

Out of 40 subjects 36 had called the two end numbers (1, 5) less often than the figure expected (40) and four subjects had called in the reverse order. (Appendix 5, pp. 339-341).

If subjects made their calls independently this difference cannot easily be attributed to chance. The probability that such a difference would occur by chance is less than 0.001.

(chi square = 25; df = 1)

This analysis only enables one to reject the null hypothesis with some confidence, i.e. the kind of patterns with respect to the different targets formed by the subjects cannot be attributed to chance. However there is no guarantee that they are due to the error of central tendency.

A counter hypothesis could be based on the following argument. Because of the group situation, one or two leaders in the group might have called a target distribution which is in close agreement with the error of central tendency. One or two such agreements could be easily attributed to chance. Others in the group may have 'followed the lead' and multiplied the two chance patterns into a significant group result.

On subjective estimation it seems unlikely that information was exchanged or obtained during the test.

The Greville correction can also be used as a kind of guide. On direct hits this correction reduced the critical ratio from 3.16 to 2.38. Even if a much stronger reduction is envisaged for the critical ratio equivalent to the chi square result obtained above, it would still remain sufficiently significant to make the above counter hypothesis to the error of central tendency hypothesis most unlikely.

The evidence that the error of central tendency is significant and fairly strong seems to support the case for GESP.

Quite apart from the theoretical probability that ESP is more likely to occur when there is no strong tendency to form response patterns (p. 102), it can be seen that if some agreement in the calls between subjects is only due to the error of central tendency, then in the application of the Greville correction, this agreement would tend to decrease the critical ratio.

Or in other words, without this error ESP would be more apparent (everything else remaining unchanged).

Analysing the target distribution in the random sequence used, it was found that there is some small agreement between this distribution and the call distribution which is presumably mainly based on the central tendency error (see Table 2).

TABLE 2

Target and mean call distribution per target figure

	Target figure				
	1	2	3	4	5
Target distribution	17	20	22	21	20
Call distribution	17.4	22.8	22.6	21.9	15.3
(group mean)					

The question which has been discussed previously (p. 103) is whether a positive deviation could be artificially high because of a particular target sequence and pre-existing tendencies to form response patterns. This problem became apparent to this author after the disguised GESF experiments had been completed. 19

It is now only possible to say that it must be assumed that the particular distribution of target numbers and the existence

19. The results were actually analysed while travelling between Australia and the USA and after arrival in the USA.

of the error of central tendency could have slightly favoured the results supporting the GESP hypothesis.

To see whether any secondary background evidence could be provided, the quarter distribution of the test session was analysed.

It is evident from Fig. 5 that this distribution follows the U curve which has been typical of a number of ESP tests (Pratt, 1949).

The difference in deviation between the first and third quarter is 48 (+ 45 in first and -3 in the third). On the basis of 80 independent runs such a deviation would yield a CR of about 2.7 (i.e. $p < 0.01$) but since the Greville correction has to be applied the critical ratio is reduced to 1.92 and the probability for this difference to occur by chance is 0.055. This difference cannot be regarded as significant.

The difference between the first and fourth quarter on the one hand, and the second and third quarter on the other, is equivalent to a deviation of 75 for 160 runs (+45, +30 for 1st and 4th and +8, -3 for 2nd and 3rd). On the basis of the Greville correction the CR is 2.1 approximately and the probability for such a difference to occur by chance is less than 0.05.

Obviously this secondary evidence is not very strong but it would tend to increase favourably the subjective estimation of background probabilities against which the main results can be judged.

Some subjects scored comparatively high in displacements of +1 and -1 but the total group did not obtain even suggestive results.

The deviations were +38 for backward displacements and -40 for forward displacements. These deviations are based on the same expectations as for direct hits. Consequently the positive deviation of 35 is somewhat too small and the numerical value of the negative deviation is slightly too large. However this has no practical bearing on the probability calculations which place the two deviations within the range which can be expected by chance fairly often.

Because of the strict control of the time sequence through interval timers it could be expected that any strong displacements due to ESP can be extracted from the data. However the slight interruptions which occur when the slide magazines are changed (every fifteen trials) could upset displacements.

The range of displacement scores is nevertheless of some interest. In both kinds of displacements these ranges are larger than for direct hits. (13-32 backwards, 15-26 direct hits, 13-30 forward).

On a speculative basis the comparatively small range for the direct hits can be regarded as a sign of a fairly stable test.

If the results are accepted at least as suggestive evidence

for GESP the questions must be considered how far was the experiment accepted by subjects in its disguised form and did any subject suspect the parapsychological nature of the experiment.

These questions cannot be answered with certainty but some background can be provided which appears to indicate that the disguise was generally accepted.

The author had not conducted any parapsychological test with any group of students previously and students were not aware of the author's interest in this field. Interest and knowledge about parapsychological matters seemed to be small among students. Such events were rarely discussed in the local press and did not feature strongly in Australian publications generally. On the other hand some students may have read Eysenck's book (1957) "Sense and nonsense in psychology" which includes a favourable introduction to ESP.

After the non-experimental pre-tests one student suggested in a private discussion with the author that ESP may have played a role in these tests. This suggestion seemed to have come about mainly through dissatisfaction with the test situation as it existed in the early pre-tests. In particular some subjects had become aware that the stimulus projected (GAP) was different from the stimuli which they were supposed to see. Because of this, the final GESP experiment was restricted to those students who had not participated in pre-tests.

In general students seemed to accept the disguised test, particularly the final version. But even if a few suspected that the purpose of the experiment was disguised there was no sign that anybody had any clear idea what the purpose might have been.

Subliminal perception was at that time a topic of some public interest in connection with advertising, and most students were anxious to contribute towards research efforts in this field. This became apparent when subjects agreed to continue in their own time with individual tests.

Nevertheless the test situation seemed to have been suitable for parapsychological processes. After the earlier non experimental pre-tests, subjects had been asked to write a short introspective report and some of these reports suggested strongly that subjects tended to rely on GESP or simply guessed rather than on visual perception: Gerald Leicester's "tendency to decide before the flash what the number would be".

The following points have emerged from this discussion:

1. The null hypothesis can be accepted or rejected depending on whether the 0.02 level is considered sufficient for the rejection of the null hypothesis.
2. Significant evidence which most likely points to the existence of the error of central tendency, was found. Agreement (in calls) between subjects due to this error would have decreased the critical ratio through the Greville correction.

3. The target sequence happened to agree to a small extent with patterns of calls which must be assumed to be due to error of central tendency, i.e. it is possible that the GESP hypothesis is somewhat favoured because of an artifact.
4. The quarter distribution - although weak in probability terms - seems to support the case for GESP by providing a more favourable background.
5. Displacement scores showed no suggestive trends but the ranges of these scores were higher than for direct hits, i.e. the comparatively small range for direct hits can be regarded as a hopeful sign of some stability.
6. It is most likely that the actual experiment was accepted by students in the disguised form. There was no indication that subjects associated the experiment with ESP. Some dissatisfaction with the earlier pre-tests did not really arise because of the disguised purpose but because of technical errors such as longer projection times which probably occurred because of "sticking" contacts. With some pre-tests it should be possible to establish a satisfactory disguised procedure similar to the one discussed here, in for instance, a different university.
7. The total result looks sufficiently promising to this author to suggest further efforts towards tests of this kind.

A GESP experiment disguised as a subliminal perception test with individual test sessions and electric shocks as reinforcements.

In the time available this experiment had to be started before any of the results of the above group test were analysed. Consequently the special problems in connection with number targets had not been considered at this stage and no changes were introduced which might have reduced this problem.

The purpose of the experiment was to see whether significant results could be obtained under similar conditions as in the group test and whether an electric shock might indirectly reinforce correct calls by 'punishing' the majority of incorrect calls.

Also an opportunity existed for three Psychology II students to conduct some sessions of this experiment. These students were left under the impression that they conducted a subliminal perception experiment. Hence it was hoped that some evidence could be obtained that would be relevant to the question whether it is feasible to attempt to develop a 'test kit' for interested scientists (p. 6).

The GESP hypothesis was to be tested by analysing the direct hits only. Other features of interest in the results were to be investigated and analysed if possible but this was not to be taken into account when the GESP hypothesis was to be tested. The experimenter expected a positive deviation but a one-tail test was not specified in advance.

Since the group test results had not been analysed at this stage the decrease of the critical ratio (of the group test results) due to the Greville correction, was not known. The author had not expected as large a decrease as actually occurred, because the conditions of the group test seemed to exclude any communication between subjects and the error of central tendency had not been considered at this stage. The randomisation procedure for targets, which was employed in the individual tests, can probably be justified if the Greville correction had led to only a small decrease in the critical ratio, but the randomisation employed now seems somewhat unsatisfactory although it ^{still} seems difficult to point to a satisfactory practical solution of this problem.

The experiments with individual subjects and electric shocks were started about two months before this writer was due to leave for the USA.

Although it would have been desirable to test the same number of 40 subjects, the individual test sessions had to be restricted in advance, to 25. This number was chosen because individual tests could only be conducted on a voluntary basis and after some preliminary inquiry it appeared unlikely that any higher number of test sessions could have been conducted in the available time. Although more than 25 students had indicated that they would be able to participate, some reduction of this number was expected if no suitable times could be arranged and it was decided in advance

that individual students should be able to participate more than once in order to complete the 25 test sessions. However only one student participated twice. Also two students who had been absent from the group test but who had not participated in the pre-experimental tests were included in this group.

The preparation of a sequence of 100 target slides was rather time consuming and the following method was adopted which allowed the experimenter to use one random sequence of slides only.

The random sequence of 100 slides was assembled in seven magazines containing 15 slides each with the exception of the seventh containing 10 only. To obtain random sequences of the seven magazines the experimenter had seven elements to vary. This number was doubled by allowing for reverse orders within each magazine.

The original order of magazine one was represented by the key number 01, 02 was the key number for the reverse order of magazine one, 03 was the key number for the original order of magazine two etc. to 14 which was the key number for the reverse order of magazine seven.

To find 50 random sequences of the orders 01 to 14 the Fischer & Yates (1948) random number tables were used. To avoid wastage of random numbers the key number 01 was selected whenever this number or any one of the following numbers occurred: 21, 41, 61,

31. The key number 02 was selected whenever 02, 22, 42, 62, 82 occurred and so on.

One two digit column of the length of the page was usually sufficient to provide a random sequence of the 14 key numbers but if a second column had to be used to complete the sequence, the next sequence was started at the top number of the next column.

To avoid the presentation of one and the same magazine twice, i.e. in the original order and the reversed order, only the first one out of a pair of key numbers was used in each sequence of 14. For instance if in a sequence 02 appeared before 01 then the first magazine was presented in reversed order only, and if 13 appeared before 14 then the seventh magazine was presented in the original order only (Appendix 6, p. 342).

50 random sequences of this kind were prepared and written out on separate scoring sheets. The random sequence for the first experimental session was drawn from these 50 sequences but this sequence was withdrawn afterwards such that the second sequence was drawn from the remaining 49 etc.

The preparation of slides was quite feasible under these conditions and consisted mainly in reversing the order of slides in some magazines. The magazines numbered from 1 to 7. Automatic projection in reverse order was not possible with the particular projector. It can often be carried out with more recent models.

The randomisation procedure is probably not entirely

satisfactory but under the circumstances it seemed the best that could be done. Subjects did not see any of their scores at the end of the session and they were also asked not to discuss the experiment with other subjects until the project was completed. However even if some subjects conveyed to others (still to be tested) that, say, stimulus 4 is the number one and stimulus 5 the number three "because I did not get a shock for these numbers", the probability that any other subject got the same sequence later is rather small. (See also discussion on additional reinforcements p. 159).

But even for future tests the preparation of a large number of slide sequences is an undesirable task, particularly because errors may occur. The answer to this problem seems to be in a machine set-up capable of selecting one of five slides according to a random number sequence which is fed in, or better still, which is self generated by the machine.

The apparatus used for the individual test was similar to the group test. The same sequence of actions was carried out by the five timers (p. 133) but the recording period (action No. 4) was reduced from 5 seconds to 3 seconds.

In addition, a panel with five press button switches numbered 1, 2, 3, 4, 5, and with five small electric globes was provided. These switches were wired so that pressing a switch resulted in the illumination of the globe at that switch as well as in the simultaneous illumination of a second globe situated on a

numbered panel in the experimenter's room. The globe at the switch served as a control to ensure that the subject would press the switch properly.

The electric shock unit was 'home' made and consisted of a 12 v' car coil' (high voltage car transformer). The low voltage input of the coil was connected to the 2v output of a power transformer which itself was wired in series with two variable resistors. The high voltage output of the car coil was connected to two finger electrodes which were interrupted by a relay switch. The shock unit was 'on' continuously but subjects received a shock only when the relay was activated by a separate circuit which, with the aid of a press button switch, could be closed from the experimenter's room. A small globe next to this switch provided a similar control as discussed above.

The strength of the electric shock could be adjusted from unnoticeable to mildly unpleasant. No opportunities for measurements existed. The shock was certainly not anywhere near the strength which may be experienced when one accidentally touches the high voltage output of the car coil when the engine is turning. To the author the shock appeared to be more of the kind which occurs if one touches, under certain conditions, the wooden base of a lamp and experiences a slight "shock & vibration".

All test sessions were conducted in two adjacent rooms in the Psychology Department at the old Domain University, Hobart (see Fig. 17, p.214). The subject's room was completely

blacked out. This room housed the 5 timers, the automatic projector, the shock unit and the panel with 5 switches. The experimenter's room housed the panel with 5 numbered lights and two switches - one to close the relay which in turn closed the electric shock circuit and one to start the sequences of 5 actions which were controlled by the five timers.

The projector with the target slides was positioned in such a way that it was practically impossible for the subject to obtain any sensory information about the slides during a test session.

With the exception of two subjects all others had participated in the previous group tests and the necessary introduction was brief.

It was indicated that the electric shock was provided to see whether they would improve their results under conditions in which they would be able to know during the test when they were right and when they were wrong.

The electrodes were then fastened to two non-adjacent fingers of one hand, this was the left hand for right handed subjects and vice versa. The shock unit was adjusted from a position of the variable resistor at which no shock was noticeable until the subjects could feel a mild unpleasant shock. There was no obvious evidence that subjects were distressed because of the shock, although a number of them felt initially somewhat anxious.

Subjects were instructed to press, according to their selection, the numbered switch button each time after the stimulus had been projected, i.e. while the room light came on. While pressing the switch button the globe next to it should light up. Subjects were informed that the time period during which they could indicate their response was reduced compared to the group experiment.

Before the actual experiment was started, subjects were given a pre-test of ten 'subliminal' stimuli similar to the pre-test in the group experiment. That is, most subjects could recognise some stimuli under the conditions of the pre-test.

At the end of this pre-test all subjects were able to carry on with the actual test without further questions. Before the experimental session was started it was indicated that the next stimuli would be much harder to recognize, but that it was again most important to indicate each time and without exception, what they thought they might have seen.

During each test session the door between the subject's and experimenter's room was closed and no communications except as specified, took place. However the experimenter had to enter the subject's room at the end of each magazine and to replace it with the next one according to the random number sequence. Those magazines which had to be presented in reverse order were prepared prior to each test session. Similarly some magazines had to be reversed to the forward order from the "reverse order" used in a previous test session.

Prior to each test session the experimenter selected at random one scoring sheet out of 50 or less (depending on how many test sessions had been conducted previously) which contained the original random sequence of 100 targets consisting of the numbers from 1 to 5 and the specific random sequence indicating the sequence and order in which the magazines were to be presented.

Magazine number 1 contained the target sequence from 1 to 15 and magazine number 2 from 16 to 30 etc. The scoring sheets were marked between 15 and 16, between 30 and 31, between 45 and 46 etc. and it was indicated by an arrow pointing up or down whether the magazine contained slides in the original order (arrow pointing down) or in reverse order. The arrows were numbered from 1 to 7 according to the random number sequence of 14 key numbers. For instance, if in such a sequence the key number 13 appeared first, then an arrow was placed in a down position on the scoring sheet starting between 90 and 91 and being labelled 1. This made it clear that the magazine No. 7 (representing the targets listed in the scoring sheet from 91-100) was to be presented first and that the presentation would occur in the original order, i.e. down the page to 100. An arrow starting between the numbers 45 and 46 and pointing upwards and being labelled "2" indicated that the magazine No. 3 was to be presented as the second magazine and that the presentation would take place in reverse order, i.e. starting at 45 and up the page to 31. It was therefore easy for

the experimenter to check whether the subject had made the correct response or not.

Since it was fairly obvious to subjects that they had a one in five chance of selecting the correct response without any additional "subliminal" information, it was decided to give them a more favourable impression of their performance than would be justified by direct hits only. It was decided to reinforce approximately half of the (-1) displacement and half of the (+1) displacements. This was carried out systematically by not giving a shock whenever a response agreed diagonally with a target in the above line of the scoring sheet. When targets were presented in the original order this was the case for post cognition and when targets were presented in the reverse order this was the case for precognition.

The students acting as experimenters were instructed accordingly.

Since the subject was not asked in this case to keep an independent record of his calls, the procedure may appear unsatisfactory. However the response "pressing a button" compared to "writing down ^a number" may be regarded as more spontaneous and it seemed therefore desirable to adopt this procedure. There was also some danger that subjects might respond to the sequential trial number.

Also since the procedure could be changed to an automatic machine recording (Keil, 1961) it seemed justified to conduct the experiment on this basis.

To obtain some idea as to how likely it was that recording errors would occur, five pre-experimental test sessions were conducted during which the subject kept a record of his calls (similarly as in the group test). The results were compared with the experimenter's recording and no disagreements were found.

Another objection to the procedure was the operation of the shock unit through the experimenter. For instance by giving electric shocks for varying lengths of time some information concerning the next target could be unconsciously transmitted. The objection was not considered to be of too great importance for the time being because it could be overcome by employing two additional interval timers (which were however, not available at the time the experiment was conducted). In such an experiment one of these timers should be started when the subject presses any one of the five buttons and should be set for a sufficient period of time to enable the experimenter to compare the subject's signal with the target and to switch an additional switch on or off (according to hit or miss). At the end of the first interval time, the second timer should close the relay circuit for a preset time. The relay will, however, only be activated if the experimenter has set his manual switch to the 'on' position. Under these conditions, which are easy to achieve with standard equipment, both the time interval between the subject's response

and shock (if given) and the duration of the shock, would be controlled independently of the experimenter.

Under the actual conditions of the experiment the experimenter simply tried to press the shock button as uniformly as possible.

Proper care was taken to prevent subjects from accidentally seeing the recording sheets (which contained the target sequence) and the psychology II students who acted as experimenters were asked to take proper care also. Some of their test sessions were partly or completely supervised by this writer.

Results.

The 24 subjects completed all their responses on 25 test sessions. A positive deviation of 46 occurred in the equivalent of 100 runs. On the basis of independent runs the corresponding critical ratio was obtained and found to be 2.3. The probability that a deviation of 46 occurred by chance is under the above conditions 0.022 or less than 0.03.

Because of the decrease in the critical ratio when the Greville correction was applied in the group test, and because of the perhaps unsatisfactory randomization procedure for this experiment (outlined p. 152), the probability estimated here must be considered to be somewhat ^{too} small. However it seems difficult to calculate a probability on the basis of the actual randomization procedure used, and the figures presented here are calculated on the

basis of independent runs. They must therefore be regarded as somewhat too favourable towards the GESP hypothesis.

TABLE 3

Summary of results from the reinforced GESP experiments with individual test sessions.

Direct hits		Displacement by -1 (postcognition)	
Deviation	+46	Deviation	+58
Standard deviation	20	Standard deviation	20
Critical ratio	2.3	Critical ratio	2.9
Probability	0.021	Probability	0.0037
Average score	5.46	Average Score	5.58
Total score	546	Total score	558
Number of runs	100	Number of runs	100

Since the displacement results are calculated on the same number of expected displacements as for direct hits, the positive deviation is actually slightly larger (Table 3).

The +1 precognitive displacements were below the expected value but did not reach a negative deviation which can be regarded as being of interest.

The total quarter distributions for direct hits and displacements by -1 tabulated from each test session (Appendix 6, pp. 343-344) are shown in Fig. 6 and Fig. 7. The differences between the quarter deviations are not suggestive.

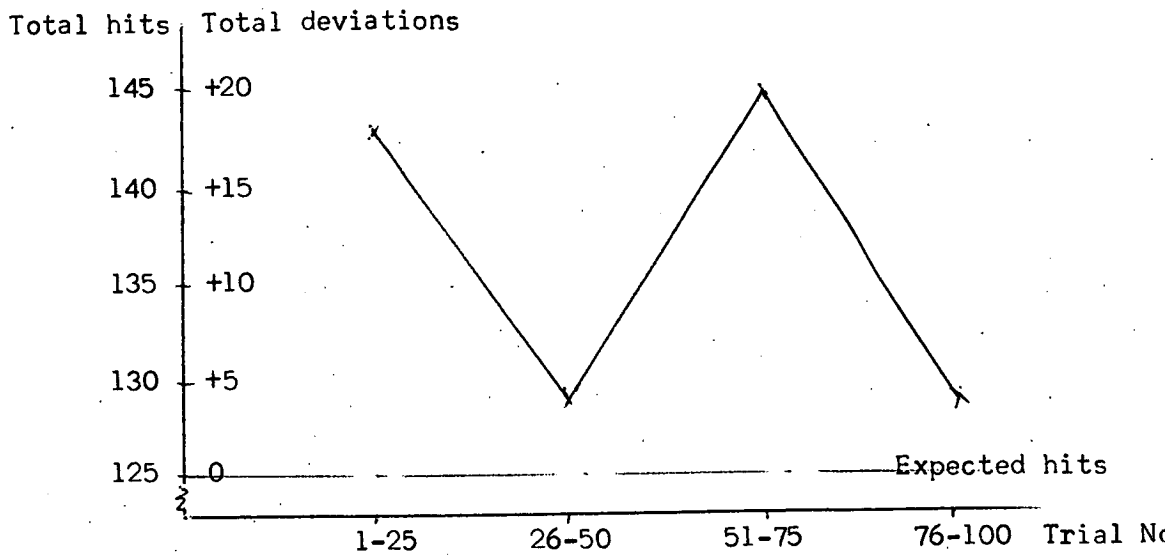


FIG. 6

Quarter distributions of direct hits in 25 individual test sessions in the disguised and reinforced GESP experiment.

24 subjects; 100 runs, total deviation = +46.

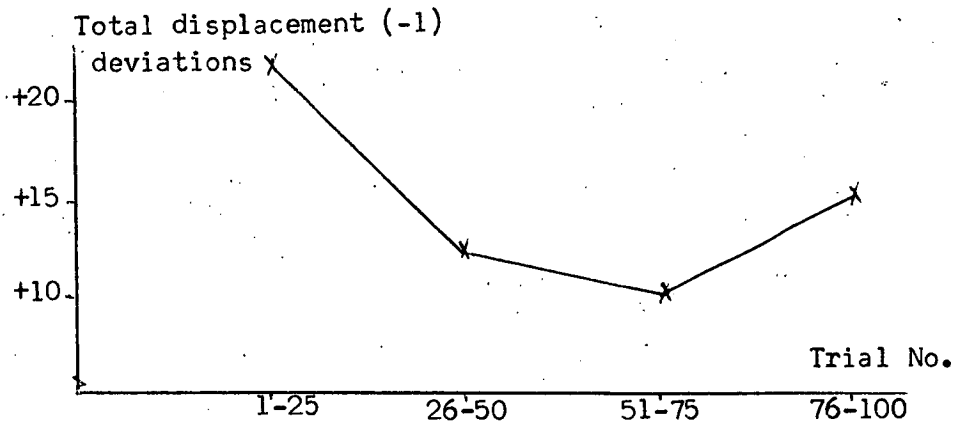


Fig. 7

Quarter distributions of (-1) displacements in 25 individual test sessions in the disguised and reinforced GESP experiment.

24 subjects; 100 runs, total displacement deviation = +58.

Out of the 25 test sessions 7 were conducted by three different experimenters and the 7 test sessions tabulated separately gave the following results (Table 4).

TABLE 4

Summary of the results of seven test sessions conducted by three experimenters

Direct hits		Displacements by -1	
Number of runs	28	Number of runs	28
Total scores	161	Total scores	150
Deviation	+21	Deviation	+10
Average score	5.75	Average score	5.36

These results analysed by themselves did not reach the 5 per cent level for direct hits or displacements but as the average scores indicate they are in good agreement with the total results.

Since the group test results indicated that subjects made their calls in agreement with the error of central tendency, the calls were analysed as indicated previously (p. 142). In these individual tests there was also a strong significant difference between the number of subjects who called the end numbers (one and five) less often than would be expected, compared to those who called these end numbers more often than would be expected.

No subject called these numbers exactly as would expected and consequently the number of subjects calling above and below should be equal or nearly so. However out of 24 subjects 21 called the numbers one and five less often than would be expected and only three called these numbers more often. The one subject who participated in two test sessions was in agreement with the former group on both occasions but the analysis was counted ⁱⁿ as one only (Appendix 6, p. 345). The difference between the two groups is significant with $p < 0.001$ ($\chi^2 = 13.5$; $df = 1$).

Discussion.

The Greville correction will not cause a large reduction in the critical ratio if the multiple calls are evenly distributed over the different target symbols called at each particular trial. Particular targets may be favoured at particular trials because of communications in the group, pre-existing call tendencies (error of central tendency) and because of the sequential position of the target. That is, if all other conditions are assumed to be equal, trial number 50 may receive more calls of the target symbol 5 than say trial number 49.

By avoiding recording booklets and by avoiding the announcement of the sequential trial numbers, it was hoped that this possibility was reduced in the individual tests.

The evidence in the group test results for the error of central tendency was not known when the individual tests were started.

Because of the time controls imposed on the group test and from subjective observation it seemed unlikely to this author that communications between subjects took place. Hence at this stage it was not expected that the Greville correction would reduce the critical ratio of the group results in a substantial way.

The randomization procedure for the individual tests (p. 152) seemed to assure much more independent runs than the group test with multiple calls. If the critical ratio of the group test had decreased only by a small amount through the Greville correction, the critical ratio of the individual results could have been considered without further correction with some justification. However a fairly large reduction of the critical ratio in the group test did occur (p. 139) and the numerical results obtained for the experiment with individual test sessions must be considered with some restraint. However a formal correction based on the limitations on the randomization procedure seems difficult since the Greville correction cannot be applied.

In spite of these difficulties the results can be viewed with some interest for the following reasons:

The individual tests have ^dsufficient similarity with the group tests to consider them as a repetition.

The (+1) and (-1) displacement deviations were positive and negative respectively as in the group test.

The comparatively high (and on face value, significant) positive deviations of the post cognitive (-1) displacements could be regarded as making sense if the decrease in the allowed response time period (from 5 to 3 seconds) is taken into account.

Similarly the change in the quarter distributions to a curve which is less in agreement with the U-shape distributions of parapsychological results in the case of direct hits (Fig. 6) could be regarded as making sense if the electric shocks are taken into account.

three
Although the test sessions conducted by Psychology II students did not provide results which can be regarded as significant by themselves, the average scores indicate that the direct hit scores were actually ^{nominally} higher during these test sessions than during all individual test sessions.

The difference between student-experimenter test sessions and the writer-experimenter test sessions is not significant and at least it can be said that there is no clear sign that the results decreased when other experimenter conducted the tests.

Since the analysis of individual results showed also significant evidence of the error of central tendency, it can be considered as confirmed that subjects will, under such conditions as described here, tend to make calls which are patterned in a way which corresponds to the error of central tendency. Since

there is no obvious sign that the electric shock made a substantial difference to the results, unless the quarter distribution of direct hits is considered to be partly due to the introduction of shocks, an attempt was made to see whether the shocks might have had undesirable effects on the tests by inducing the subject to make calls depending on whether the symbol called last had been followed (or not followed) by an electric shock.

For this purpose all symbols which were called and repeated once were counted and put into two classes (per subject).

1. Repetition after the previous symbol called had received a shock.
2. Repetition after the previous symbol called had received no shock.

The number counted under 1. was expressed as the numerator of a fraction where the denominator expressed the number of shocks which a subject received during one test session.

The number counted under 2. was expressed as the numerator of a fraction where the denominator expressed the number of trials during which a subject received no shocks during one test session.

By chance there should be similar numbers of subjects with fraction 'one' smaller than fraction 'two' compared to the number of subjects with fraction 'one' larger than fraction 'two' and perhaps some subjects with equal fractions.

For the one subject who participated in two test sessions the two type 'one' fractions were added and compared with the two type 'two' fractions. That is, the two sessions were combined and counted only once in a chi square analysis. One subject gave no repetitions at all. For all other subjects the two established fractions were not equal (Appendix 6, p. 346).

Since it was difficult to predict what effect the shocks should have, the subject with no repetitions was excluded. Otherwise this subject might have been included in the group disfavoring the predicted results. Hence 23 subjects were left for analysis. Of these 23 subjects 15 gave more repetitions than would be expected after they had received a shock previously, and 8 gave more repetitions than would be expected after they had received no shock previously.

The calculated chi square is 2.1 approximately and the probability for this difference to occur by chance is $0.1 < p < 0.2$. This difference is not significant.

Since this test is not directly concerned with parapsychological events it may be justified to discuss the trend represented by the above figures. If this trend is meaningful, then the shocks which were received by the subjects after calling certain symbols, induced the subjects to call those symbols again more often than would be expected on the basis of their general tendency to repeat symbols.

It is difficult to say whether one should expect the opposite or not. On the one hand one may expect that subjects avoid those symbols which were associated with shocks previously, on the other hand subjects might more or less consciously have expected that because certain symbols had been shocked previously they are not so likely to be shocked again.

Such considerations may still have an influence on a subject's choice even if the subject is aware (on rational grounds), that the probability of a particular target symbol occurring is independent of previous target symbols (J. Cohen, 1960).

Since the actual shocks were rather weak (p. 155) they may not necessarily have induced the subject to avoid them but may instead have directed more attention to those symbols which received shocks. From the limited evidence presented it cannot be concluded that subjects formed strong patterns because of the shocks. Even if weak patterns were formed, from the results presented here it would appear that these patterns should not have disturbed the ESP output in a marked degree, but the picture may change considerably if for instance the strength of the shock is increased.

If the total results of the individual tests are regarded as relevant to parapsychological events, the disguise of the test should be considered again.

There was no indication that subjects did not accept the test as it was presented to them. The subjective evidence is,

in this case, better, since the experimenter had an opportunity to exchange a few remarks with most subjects about the test at the end of the test sessions.

From these discussions it became clear that subjects favoured generally the individual tests in which they felt more confident because they had received less shocks than they expected.

It is less certain how far the students acting as experimenters accepted the experiment entirely in its disguised form. They had heard or read about parapsychological experiments and some considered the possibility that parapsychological processes might play a part in the experiments.

The main methodological doubt about this experiment seems to be based on the randomization procedure. To overcome this problem it would be necessary either to prepare carefully a large amount of test material or to design a suitable apparatus.

A third possibility may be to rely on the target list viewed by the experimenter only, and to present no target slides at all. This does not require any major technical changes. The target slides could be left out or replaced by one and the same non-target symbol. The action sequence of the five timers may be simplified.

Other weaknesses in the procedure discussed previously (pp. 159-161) seem in comparison less serious, particularly since they can be overcome by the application of standard equipment.

The PK apparatus

To explore as systematically as possible the research possibilities outlined (p. 111, Appendix 3, pp. 310-330), a substantial proportion of the total effort was devoted to the design and the construction of a versatile PK apparatus.

The recognition by parapsychologists that fraud must be accepted as a possible counterhypothesis (Rhine & Pratt, 1961) and that steps must be taken by parapsychologists to exclude this counterhypothesis as effectively as possible, has led to a research situation which is not entirely satisfactory.

It had led to the two-experimenter approach which has a number of advantages but it has not really changed the attitude of the sceptic (e.g. G. Price, 1955; Hensel, 1961a, 1961b) and obviously if every single parapsychologist is suspect of fraud, then the pairing of such suspicious individuals does not guarantee the exclusion of fraud, although it is likely that this possibility is further reduced.

A two experimenter approach may, in the best case, result in the fruitful co-operation between two parapsychologists, it may help to eliminate technical errors and it may help to penetrate to a deeper level of the problem at hand. If the basic crucial problems in parapsychology are complex - and this is likely (Murphy, 1958) - then a team of specialists from various fields may be necessary to solve these problems. But in this sense

the advantages are similar in most other research fields. In parapsychology however, significant results - obtained in the more than one experimenter setting - may be expected to be repeated more easily. So far there is no clear evidence that this is the case.

Research work in a new field is perhaps more than elsewhere individualistic. Parapsychologists are at present working in comparatively small groups and often alone with no co-worker easily at hand over a considerable distance.

The anti-fraud suggestions made by G.R.Price (1955, 1956) are perhaps unrealistic for the average experiment but it seems desirable nevertheless to arrive at a design which would allow one experimenter to proceed under conditions which are equal or perhaps even better than those in the two experimenter approach, at least as far as the attempts to exclude fraud are concerned.

The conditions under which the R. A. McConnell experiments (McConnell, Snowden & Powell, 1955) were carried out could perhaps have been adapted in this way. The photographic record could have been sealed, precrossed and scored by an independent party who otherwise could have remained dissociated with the research.

The present writer had attempted earlier to adapt a tape recorder as an automatic recording device. It was found that pen markings did not work satisfactorily, i.e. some ambiguous markings occurred but above all the visual scoring of marks on a tape is quite tedious and time consuming.

As a first step towards the PK apparatus an automatic impulse and scoring device was designed and constructed. In this case a standard electric adding machine (Addo X) which prints numbers on a paper roll, was adapted for automatic operation. Since the PK apparatus was planned to have only two kinds of scores (i.e. not six as in the case of dice) only two kinds of numerical responses were needed. The numbers 1 and 100 were selected and by adding each set of ten responses the sum (printed by the machine conveniently in red) of say 307, stated clearly that the last ten responses consisted of three responses of kind 100 and seven responses of kind 1. Since each entry was printed it was also possible to consider single responses or any patterning of these responses, i.e. the 307 may be preceded by the following sequence (Table 5).

Table 5

Example of a set of 10 recorded responses of two different kinds.

100
1
1
1
100
1
100
1
1
1
<hr/> 307

This recording system is more efficient than the tape marking and a long sequence of results can be analysed conveniently.

Ambiguities are avoided provided that certain kinds of electro-mechanical errors are eliminated and that the correct operation of the device can be checked very easily, e.g. by checking whether 3 and 7 from 307 add to 10.

The construction of a fairly simple adaptor to the 'Addo X' (see pp. 199-201), led, it is hoped, to a more general aid in scoring parapsychological and psychological responses automatically (Keil, 1961). Most, if not all of the possibilities outlined in this paper (Appendix 6, p. 347), can probably be achieved with different kinds of apparatus. But the apparently simple and inexpensive construction (the necessary parts may be purchased for between A£5 and A£20 = US \$10 - \$40) may make automatic scoring feasible where it was previously considered to be too expensive. This seems to have some particular relevance to the field of parapsychology. Scoring errors which have been criticized by Kennedy (1952) can be eliminated without too much double checking.

By increasing the number-responses to be scored (Appendix 6, p. 347) the 'Addo X' adaptor could be used for such tests as the GESP test for individual sessions (p. 159).

The possibility of fraud can be excluded or at least reduced by an independent party who makes a copy of summary of the numbers recorded before they are used for further analysis.

Since the adaptor works through electric impulses which come from the PK apparatus, it is in principle possible and practical to use two adding machines and thus to obtain two identical records, one of which could remain sealed throughout the experiment.

Of course the fraudulent change of records is only excluded in as far as it is made impossible to change the impulses from the PK apparatus. But it is likely that some agreement could be reached on this point as G. R. Price's papers (1955, 1956) suggested.

Perhaps the importance of excluding fraud is here unduly emphasised. But it should only be seen as the last consideration in the evaluation of the usefulness of an apparatus. It is hoped that recording errors are virtually eliminated but there was also one major consideration which led to this particular design.

The earlier critical suggestions (Brown, 1953) and Girden's more recent one (1962a), that parapsychological results should be checked against the empirical random results of whatever randomisation device was used, led first of all to demonstrations that the results cannot be explained away through such comparisons and increased confidence in the "internal checking designs". This writer is in full agreement with Murphy (1962) on the adequacy of the internal check, but it seems that there exists at least one research possibility for which an external empirical random distribution is necessary.

As mentioned previously (p. 3) there is now a good deal of evidence that parapsychological results occur but in the opposite direction to the subject's overt wishes. Displacement may be seen as a deviation from the "correct direction".

Strictly speaking it is not necessary to demonstrate a particular "direction" as long as it can be shown that the distribution which occurs when a subject tries to do something (although the subject should usually be instructed to wish for a specific result for motivational reasons), is significantly different from the distribution which is automatically generated and recorded by the apparatus.

From earlier observations it appears that PK results may be particularly strong (p. 242) if the subject is emotionally
20
involved.

It was also recognized, as was pointed out again recently (Murphy, 1962), that the experimenter may unintentionally influence the results and that either he or the subject or both, ⁱⁿ may influence the results during periods which no PK effort is planned according to the experimental design.

20. In this case the GESP hypothesis (instead of the PK hypothesis) is unlikely because of the long time changes which were of the order of 5 to 10 seconds compared to the usual time changes which were of the order of 0.3 to 0.7 seconds.

Because of this it seemed less promising to compare in short succession a few PK trials with a few automatic trials.

It was therefore planned to design and build a PK apparatus and recording device capable of running and recording continuously and automatically without any attendance for 6 to 12 or more hours. The 'Adde X' adaptor which was built first scored a continuous sequence of 1200 responses without error. However at this stage the impulses were provided through contacts made by a constant speed motor.

Because of the necessity to provide electric impulses the traditional PK objects, i.e. dice, were considered less suitable. Spinning a coin (Theouless, 1945) could be adapted more easily as far as two different responses are concerned by providing one dull and one reflecting side which could with suitable illumination and a photoelectric cell, provide the necessary impulses for the 'Adde X' adaptor. However the present writer found the design and construction of a mechanical apparatus which would have to pick up the coin and spin it, too complicated to be carried out under very limited workshop facilities, and consequently the path of a steel ball (from a ball bearing) was used to provide an experimental setting in which PK may occur.

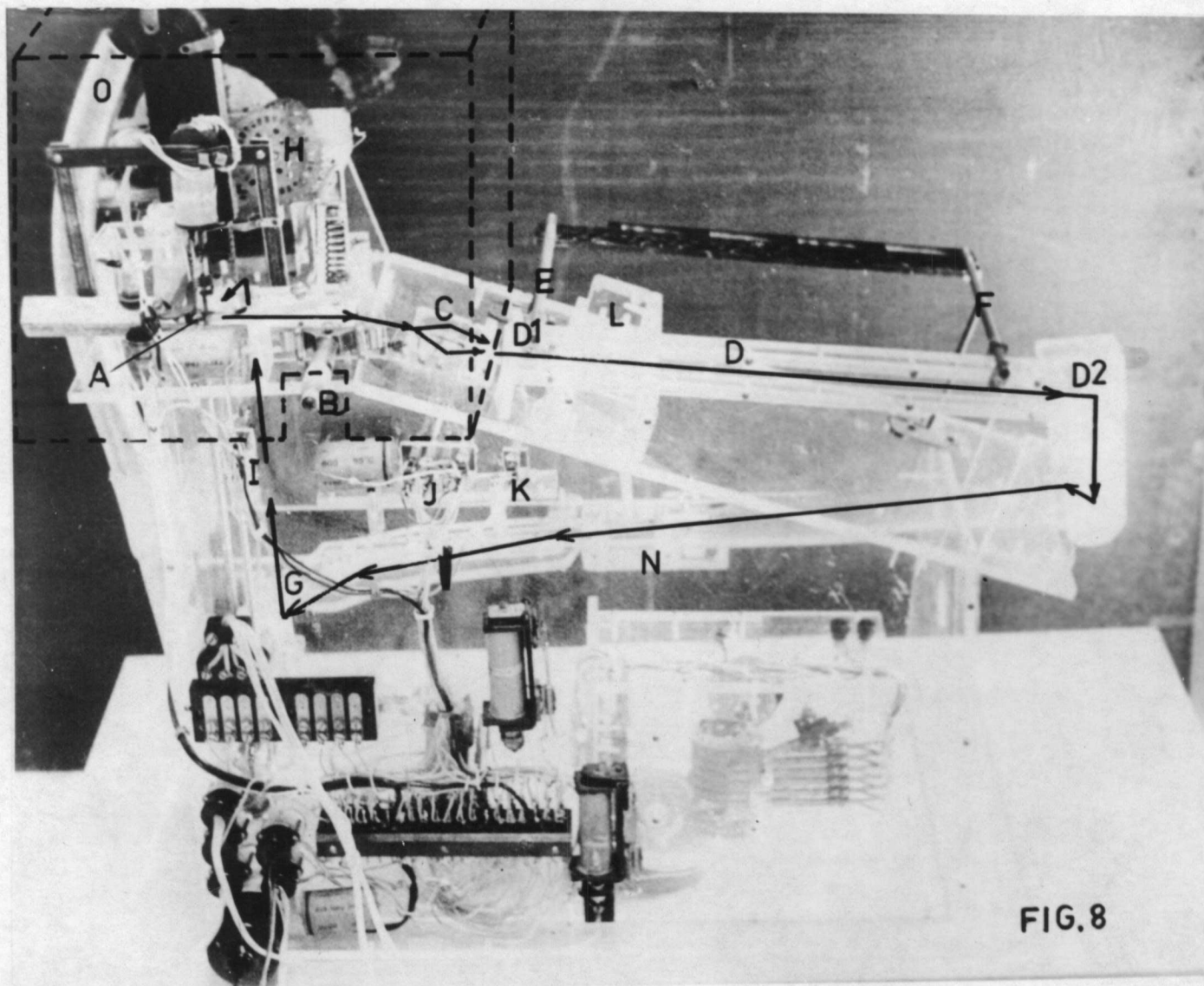


FIG. 8

Fig. 8. PK apparatus

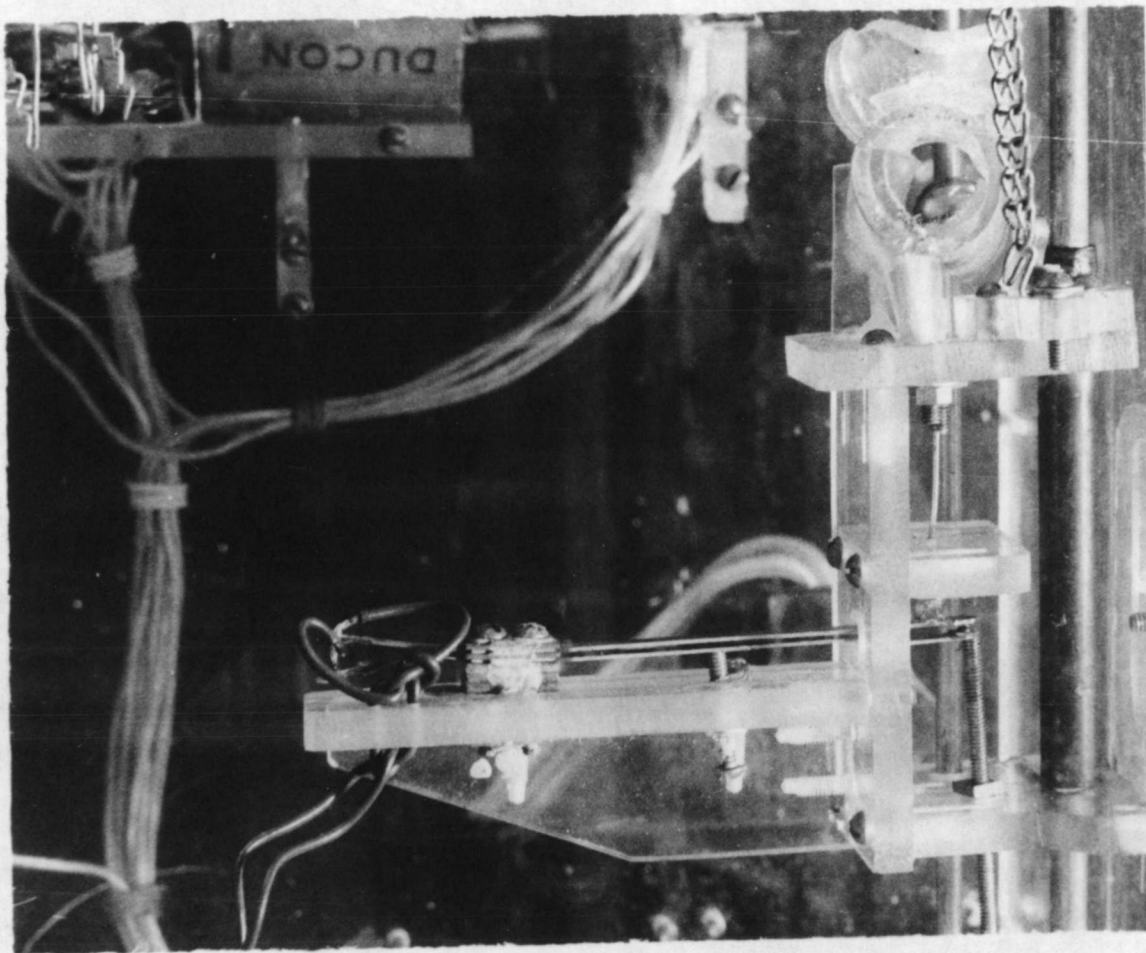


Fig. 9 Sledge to move ball back to the releasing position. The ball is just visible (upper right hand corner) in the plastic tube which is part of the sledge.

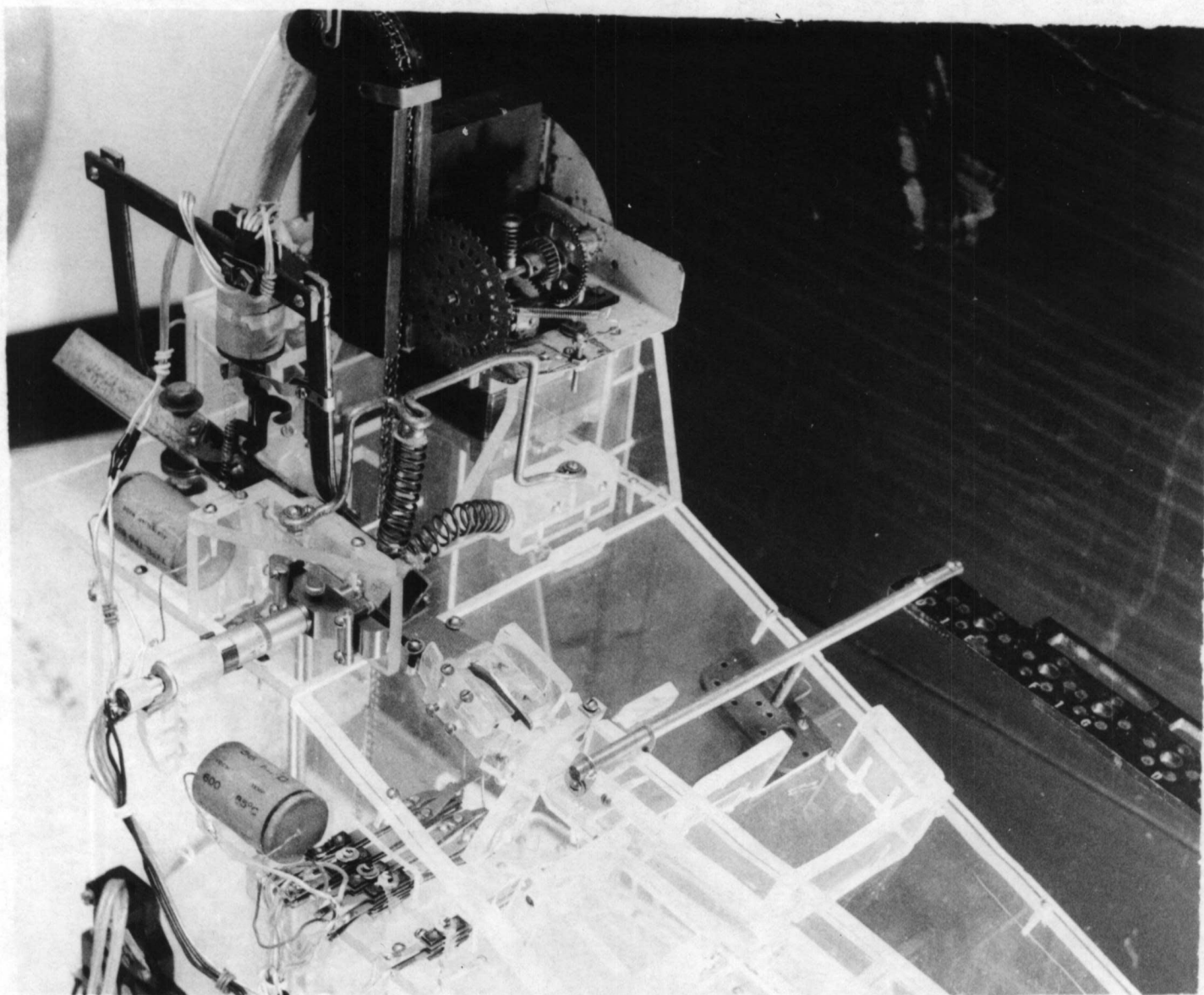


Fig. 10

PK apparatus

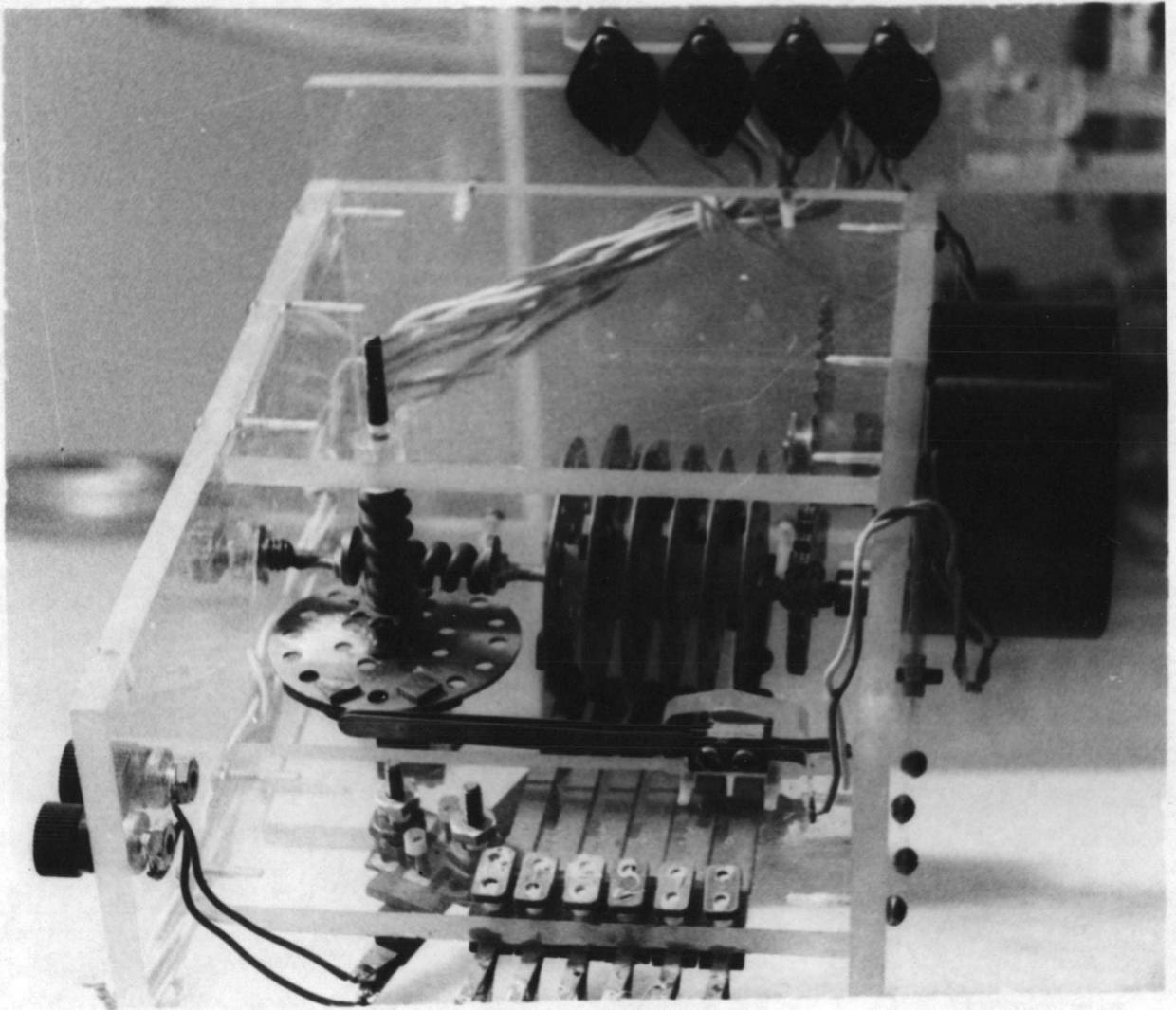


Fig. 11 Control unit

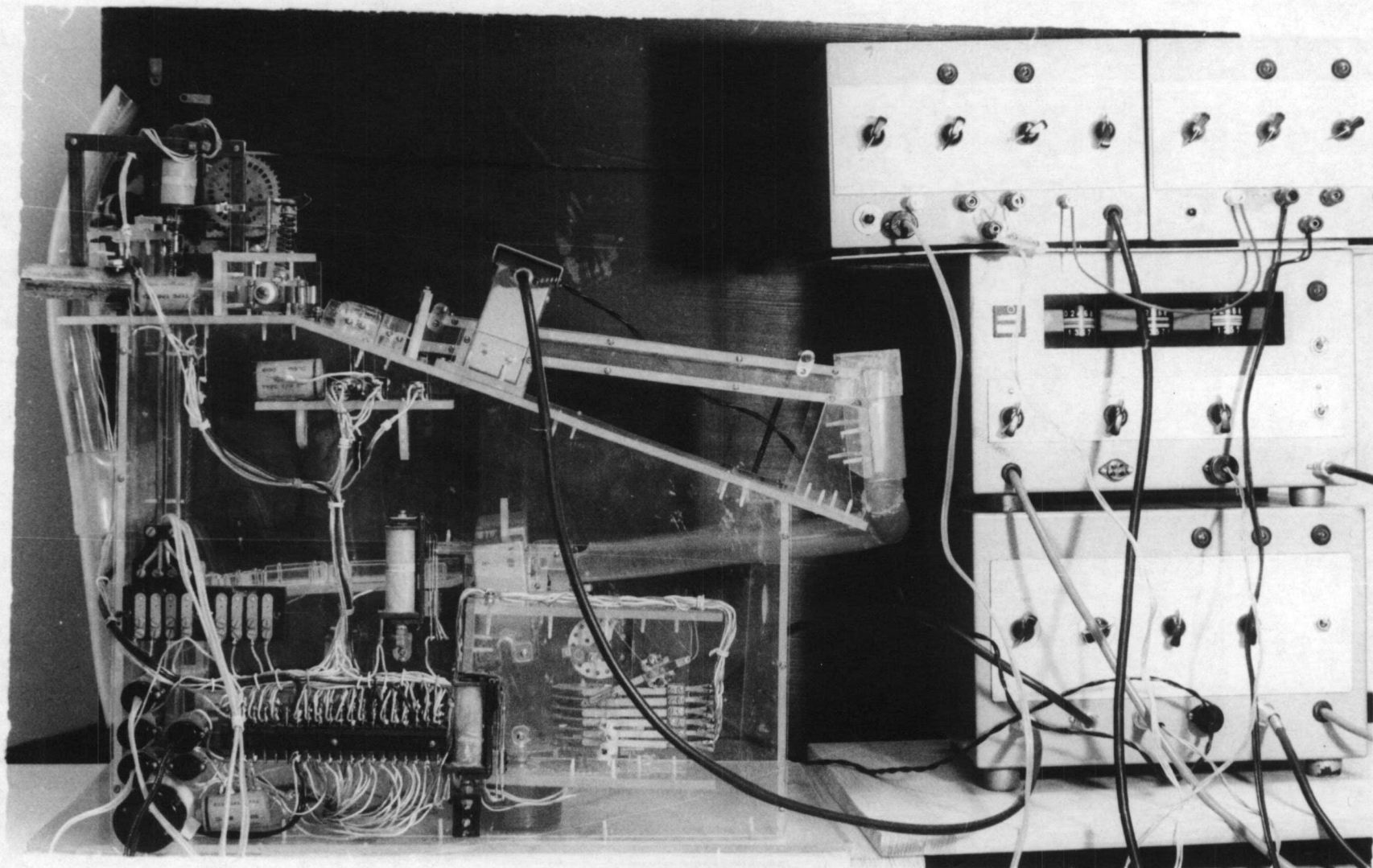


Fig. 12 PK apparatus with two photoelectric cell and light units connected to the oscillator (bottom). The decatron counter (middle) supports the two interval timers (top).

The ball is electromechanically released at (A) see Figs. 8 and 10. It then rolls down a slope until it comes to a hardened steel wedge at (c) where the ball rolls past this wedge either through right or left hand channel. These two channels are joined again at (D1) where the ball enters a chute which is part of the balance with fulcrum at (E) and (F). The ball then rolls to the lowest point (G) which is part of a sledge (see Fig. 9). The weight of the ball closes the contacts (Fig. 9) which electromagnetically engages gear (H) see Fig. 8, in the chain (Figs. 9 and 10). When the sledge has moved up far enough the ball rolls down into the releasing position (A). As soon as the ball leaves the sledge the contact is broken, the gear is disengaged from the chain (through spring loading) and the sledge moves down again.

Fig. 11 shows the control unit with a one minute constant speed motor geared to produce two trials a minute.²¹ The six disks on one axis close contacts for various proportions of 30 seconds. The separate disk completes one revolution every 10 trials (or every 5 minutes) and closes the contact which engages the "sum" mechanism of the 'Addo X' adaptor. The other six contacts have the following purpose (in sequential order) :

21. This unit is housed in the PK apparatus (Fig. 8) but it can be set up separately which is desirable for initial adjustments.

1. The number 1 is pressed into the 'Addo X' adaptor.
2. The ball is released at (A).
3. An electromechanical digital counter receives an impulse (counts one number).

(If the ball proceeded through the left channel it would by its weight close two contacts which press two zeros into the 'Addo X' adaptor).

4. The 'Add' button is pressed on the 'Addo X' adaptor. Thus either the number 1 or the number 100 is printed.
5. One of the two contacts wired in series for the 'sum' button of the 'Addo X' adaptor is closed. If after 10 revolutions the separate disk also makes contact, the sum of the ten previous trials is taken and printed. The series of two contacts is necessary to ensure a correct time sequence of operations.
6. The motor turning gear (H) is switched on for a sufficient period to bring the sledge up. This period is somewhat longer than the strictly necessary time. The gear (H) can turn freely when it does not engage in the chain and it is disengaged as soon as the ball leaves the sledge. The motor is turned off before the ball is released at (A). That is no mechanical movement is going on in the apparatus at the time of the release.

During one trial the ball can make the following contacts (depending on what channel is taken):

The ball closes (through its weight) contacts at (I). These contacts lead to a count on an electromechanical counter which should tally with a similar counter which is operated through the control unit. At (J) there are four contacts, two for each side of the channel. For the left channel two contacts result in one zero each to be printed through the 'Addo X'. The right channel path results in the switching on of one electronic timer for a pre-set time and in one count on a third electromechanical counter. The contact (K) is closed for the time the ball runs through the chute between D1 and D2. During this time a relay is closed which can close three independent circuits simultaneously. Finally the ball closes the sledge contact when it reaches point (G).

In conjunction with the apparatus two powerpoints P1 and P2 are used which supply electricity to any stimulus device such as an electric shock unit or a slide projector. However only one of the two powerpoints will be 'on' during each trial depending on whether the timer is switched on when the ball moves to the right channel. Either of these powerpoints is only 'on' during the time the ball runs into the chute between D1 and D2, i.e. when contact (K) is closed.

If the ball enters the left channel, the timer will remain 'off' during this trial and P1 is on (when K is closed) and P2 off.

If the ball enters the right hand channel the timer will switch P1 'off' and P2 'on' for 15 seconds. However since the 'on' powerpoint is in series with the relay contacts which are operated through contact (K), both P1 and P2 will only be 'on' when (K) is closed.

At the end of the 15 seconds period the timer will restore the original setting : P1 'on' and P2 'off'.

The arrangement makes it possible to present to the subject either of two electrically operates stimuli which start within a fraction of one second after the ball has reached (C). Which stimulus is presented depends on which lane the ball entered after reaching (C).

By providing one positive (or overtly desirable) and one negative (or overtly undesirable) stimulus to the subjects, a situation is provided where correct PK responses can be reinforced.

The difficulty of selecting suitable stimuli will be discussed later (pp. 236-240) but the main stimuli used were colour slides as positive and electric shocks as negative stimuli.

To keep up interest in the positive stimuli the slides were changed up to 60 times in an automatic Agfa projector.

Two magazines with 30 slides each were joined together and proceeded through the projector satisfactorily if some supports

were added on which the extreme ends of the magazines could move. However after 30 slides had been projected one blank would appear before the 31st slide was projected.

The contacts necessary to change the slides were made through a second interval timer (see fig. 12) which was switched on for 0.5 seconds each time the first timer switched itself off at the end of 15 seconds. An impulse at the end of the set period of the first timer triggered the second timer.

The actual time either of the two stimuli was on was of the order of 1 second.

By including the balance which makes a small movement through which (K) is closed while the ball moves from D1 to D2 a second possibility is provided through which PK may be investigated.

Obviously subjects may be expected to wish for short times when the shock stimulus is 'on' and for long times when a desirable slide stimulus is 'on'.

One would have to expect different times for the D1 to D2 run (see Figs. 8 & 10) of the ball depending on the channel the ball had entered previously, i.e. the initial velocity and the direction of the ball is likely to be different on account of the different channels. Nevertheless any changes due to PK may be detected by comparing the left channel D1 to D2 times of the automatic trials (no subjects present), with the left channel D1 to D2 times when the subjects received a negative stimulus and similarly when the subject received a positive stimulus. (Similar comparisons may be made for the right channel).

To record the D1 to D2 times the relay which is operated through the (K) contacts could, simultaneously with the stimulus for the subject, switch 'on' a tape recorder for these D1 to D2 times. Other contacts could be used to mark the tape at the beginning and end of these periods as well as to indicate (e.g. by two dots and one dot) whether the ball passed through the left or right channel.

As indicated (p. 173), the analysis of such tape markings is difficult. Nevertheless the availability of commercial tape-slide synchronisers may provide a practical possibility.²²

Some future attempts in this direction are desirable because the record of numbers does not include any information on the D1 to D2 times.

Fig. 12 shows one further arrangement for recording times through photoelectric relays situated at (L) and (N) (see Fig. 8). The relay at (L) starts an ultra stable oscillator which is connected with a decatron counter which is stopped again by the second relay at (N). The time measurements do not strictly coincide with the D1 to D2 measurements but it may be expected that the D2 to (N) times do not interfere with any measured PK changes which may occur between D1 and D2 and which indeed may continue somewhat past D2.

22. These synchronisers provide the means of putting a sound signal on tape which on playback later can be used to activate a relay.

The decaatron counter probably provides an unambiguous reading which, however, has to be observed by an experimenter (or a camera) and which has to be recorded. Continuous readings of this kind were therefore not possible in the way it was done for the left right distribution of the ball.

The decaatron counter set-up, triggered by photoelectric relays, was successfully tested but in the time available no experiments were attempted. For future experiments a combination of the tape recording with the aid of the tape-slide synchroniser and the decaatron counter seems desirable. The tape should be kept as a permanent record which can be checked against the decaatron measurements at a later stage in a semi-automatic way (see also pp. 241-243).

The circuits of the PK apparatus operated on 240 volts AC and on 12 volts DC. The low DC current was used to avoid erratic operation of relays and to reduce the wear of silver contacts through sparks. To reduce the sparks still further condensers were wired parallel to the contacts.

The apparatus was mounted in a corner room in the old Tasmanian University (Domain) by bolting steel brackets into the stone walls (Fig. 17, p. 214).

The 50/50 distribution of left-right paths of the ball could be approximated through the depth micrometer (B) see Figs. 8 & 10, which moves the steel channel through which the ball runs after being released at (A).

A hardened piece of steel file with a rough criss-cross pattern (see Fig.10) may have provided more than one position at which the right or left path of the ball is decided. Evidence of more than one decision point was obtained when it was found that a 50/50 distribution could be approximated at two distinct micrometer settings with a one-sided distribution between the two settings.

The top of the apparatus was covered with a perspex box which had been removed for the photos included here, but which is drawn in dotted outline in Fig. 8.

To come somewhat closer to the 'ideal conditions' for the description of an experimental set up (Murphy, 1962), some further details about the apparatus will be introduced here and construction and operation difficulties will be discussed (pp. 192-201 and 204-207). However, it is hoped that the previous description is sufficient to visualise the experimental procedures carried out with this apparatus.

Additional details about the PK apparatus

The material which was mainly used for the construction was $\frac{1}{4}$ " perspex. To allow for transportation the perspex pieces were only bolted together by $\frac{1}{8}$ " bolts (see Fig. 11) except for some small parts which were glued.

On a firm base (steel brackets bolted into stone walls), the apparatus showed no clear signs of a systematic drift in the left-right distribution of the ball, (see Table 6, p. 203).

If necessary, perspex could be glued into a much more rigid frame, particularly since glueing of this material with the proper cement amounts to a welding process.

The release mechanism of the ball at (A) was bolted onto a square steel channel. At the upper end this steel channel was fastened fairly securely with a $\frac{1}{4}$ " bolt to the perspex base. At the lower end of the channel a small ball (from a ball bearing) was soldered at the bottom edge of the steel and a larger ball was soldered at each side and at the top of the steel channel. The steel spring above the channel presses the small ball to a flat piece of polished and hardened steel fastened to the perspex base and the steel spring at the side presses the steel ball on the other side of the channel against the flat spindle of the depth micrometer (Fig. 13).

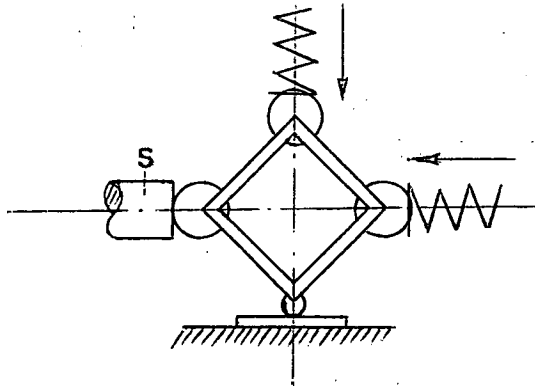


FIG. 13

Simplified drawing of lower end of the release channel.

Not to scale. S: spindle of depth micrometer.

It appears that this is a satisfactory arrangement for adjusting the release channel to a position where a distribution close to 50/50 will occur.

The necessary movement for adjustment is rather small and of the order of 1 mm. at the lower end. Within such limits the movement does not interfere with the steel pin which comes up through a hole in the perspex base and through a hole in the steel channel to enable the ball to close contact (1) shortly after it has been released.

The actual release mechanism consists of a steel pin against which the ball rests and which moves up electromagnetically to release the ball.

The return of the ball to the release position with the sledge (Fig. 9) provides more difficulties which had to be overcome.

A simple improvement which was awkward to introduce later but which would minimize adjustment difficulties is to allow for more space once the sledge has reached the position from which the ball rolls into the releasing channel. Even without spare space the sledge and the return mechanism worked satisfactorily over thousands of trials but occasional failures occurred which could probably be avoided through additional space at the top.

The upward movement of the sledge was achieved by engaging gear (H) into a chain which is fastened to the sledge. The other end of the chain is weighted and slightly spring loaded and moves inside the plastic tube (O) (Fig. 8).

The downward movement of the sledge occurs through gravity only. In order to stop the sledge at the lower position gradually, the spring loading of the other end of the chain reaches a maximum for the lowest position. The final lower resting point of the sledge is on top of a steel spring which also helps to bring the sledge to a less sudden stop.

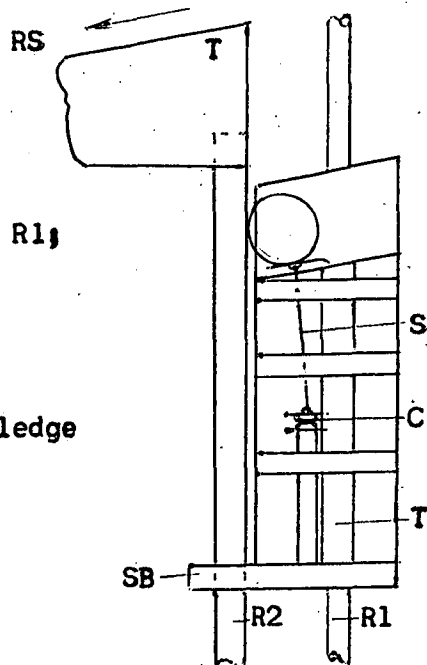
The path of the sledge is controlled by two parallel steel rods. One on the right hand side of Fig. 9 is clearly visible, as well as a brass tube which slides on the steel rod. A small section of the second steel rod is visible behind the ball which can be partly seen inside the sloped plastic tube which is part of the sledge. A hole cut into the lowest perspex base plate of the sledge allows the sledge to move along this second rod in a particular position only. This second rod also prevents the ball from rolling out of the plastic tube until it reaches (T) (Fig. 14).

FIG. 14

Simplified drawing of sledge.

Not to scale. T: top; RS: release channel; T: tube which glides on rod R1; SB: sledge base; S: steel pin; C: contacts.

This is a view from the LHS of the sledge compared to FIG. 9.



To avoid vibration of the contacts while the sledge is moving, two adjustable bolts are moved against the lower contact (Fig.9).

A tape recorder motor was used to provide constant speed for the chain and sledge movement. The vertical axis of the motor with a worm gear mounted on it is visible in Fig. 10. This worm gear is continuously engaged in another gear which is mounted on the same axis as gear (H) (Fig. 8). This axis rotates in a U-steel bracket which itself can rotate slightly on a vertical axis which is situated approximately underneath the gear which is driven by the worm gear of the motor (Fig. 15).

The movement of the bracket and gears and the engaging of gear (H) onto the chain is carried out electromagnetically. The bracket is returned against a stop (Fig. 15) by spring loading. The chain moves in a groove which was cut into a piece of hard wood. The back of this groove was lined with thin steel and a slot was cut into the steel to allow the teeth of the gear to engage fully into the chain. No difficulties in the movements of the chain were experienced.

Once the sledge reaches the position from which the ball rolls by its own weight into the releasing channel, the sledge can only move another 5 to 7 mm. before touching other parts of the apparatus.

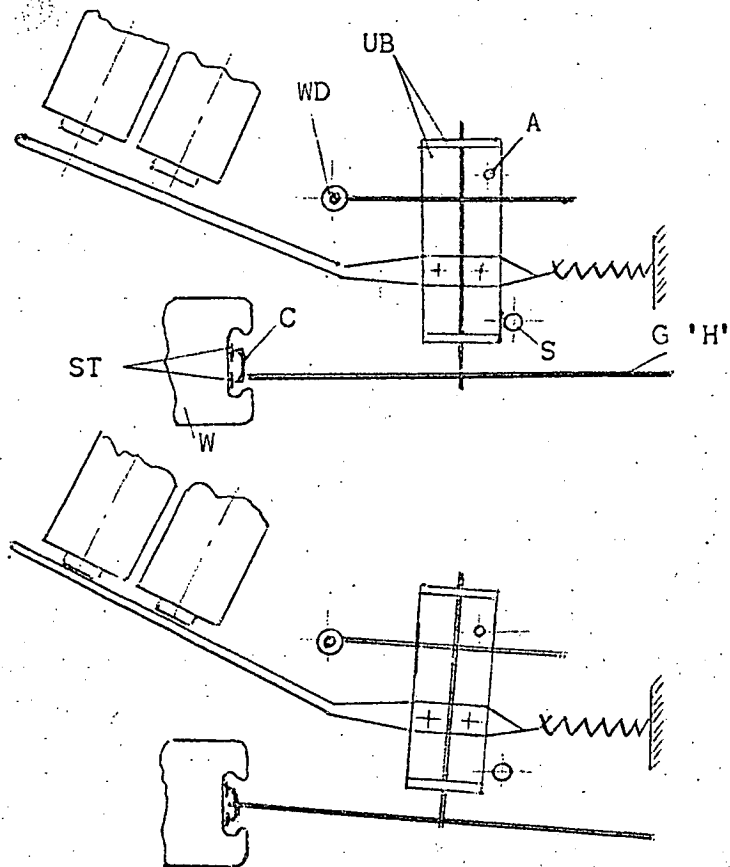


FIG. 15

Simplified drawings of gear and chain arrangement to move sledge. Not to scale.

UB: U-bracket; WD: worm drive; G: gear; A: axis around which U-bracket can move; S: stop; C: chain;

ST: steel track with cut in slot; W: wood mounting;

G 'H': gear H.

Top: disengaged position.

Bottom: engaged position.

If the sledge was allowed to move on a good deal further, say about 150 to 200 mm., the timing of the disengagement of gear (H) would be less critical and could be monitored from the control unit rather than from the ball.

Attempts to use the control unit for this purpose were not successful within the narrow margin of movement.

The sledge contacts, closed by the ball, tended to wear out through sparks more quickly than other contacts, and this seemed to be due to the movement of the sledge which still tended to result in some vibration of the contacts even when this was minimized through the two bolts mentioned above.

All contacts were silver contacts from obsolete P&G equipment. The contacts worked satisfactorily for thousands of trials but from a certain point onwards occasional errors occurred and some contacts were replaced.

The cost of the material for the apparatus was very low indeed. It is doubtful whether new silver contacts would have lasted any longer than the ex P&G contacts which were in new condition and perhaps not used previously at all.

Contact wear was further reduced by introducing reduced loads on the circuits which showed most wear. This was done by introducing additional relays which in turn closed contacts for the necessary final loads of, for instance, the electromagnetic coils which move gear (H) against the chain.

even
 However for continuous and unattended operation one error in 1000 trials can be troublesome, and with more finance available it might be worth while to enquire for highly reliable contacts.

Mercury contacts are probably more reliable than ordinary silver contacts but it would be more difficult to set up these contacts which require a fair amount of movement. The rolling ball provided, through its weight and the movement of the pins, a rather limited mechanical movement only. Nevertheless some preliminary investigation into the use of these contacts would be desirable.

The diameter of the steel ball is $\frac{1}{8}$ ". Smaller balls would reduce the weight and thereby introduce difficulties inclosing contacts by weight. Larger diameters would increase the danger that the movement of the ball might change the position of some parts of the apparatus which could lead to a systematic change in the 50/50 distribution. Of course some changes in the diameter of the ball might actually improve the whole set-up since it is not certain that the $\frac{1}{8}$ " size is the best choice but substantial alterations seem undesirable.

The attachment to the adding machine was made from two bakelite plates $\frac{1}{8}$ " thick, (also ex P&G) which are fastened to the machine in such a way that the whole attachment can be removed easily, i.e. the machine may be used for ordinary office purposes (Fig. 16).

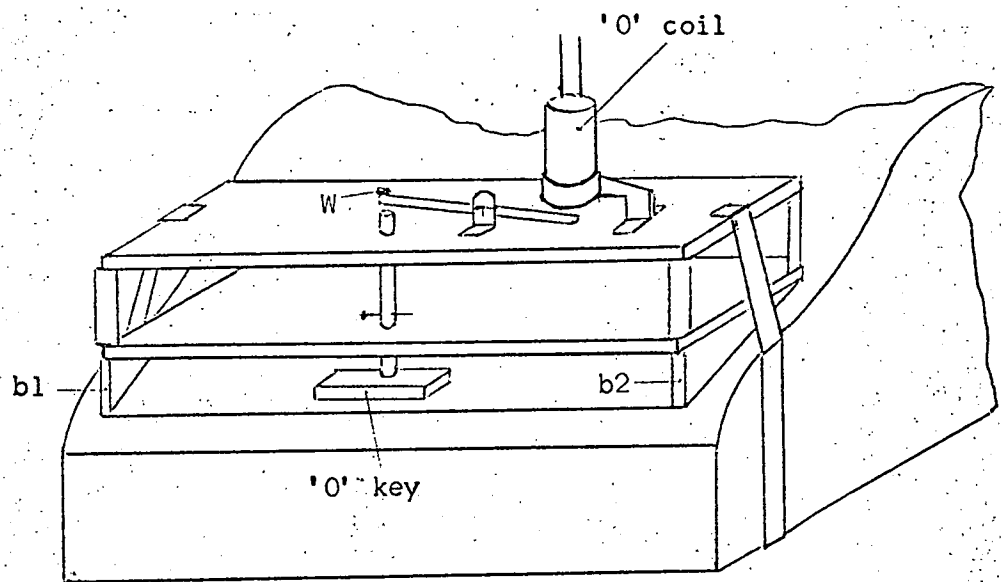


FIG. 16

Simplified sketch of 'Addo X' recording attachment.

W: weight; b1 and b2: bolts which hold the attachment in position.

Bolts b1 and b2 fit into the corners of a recess of the adding machine and one simple elastic band will hold the whole attachment in correct position. For the sake of clarity Fig. 16 shows one electromagnetic coil only, i.e. the 'O' coil which presses the 'O' key of the machine.

The particular adding machine (Addo X) has only one key per digit from zero to nine. To add 100 the 'one' key has to be pressed once and the 'zero' key twice (Appendix 6, p. 354).

The actual attachment has three further coils, one for entering the number '1' one for 'add' and one for 'sum'. The operation of this unit did not provide substantial difficulties however it happened that the coils were comparatively weak and it was necessary to add some weight to the lever at (W) to overcome some of the spring loadings of the keys (Fig. 16).

Tests for Randomisation.

To test whether the machine produced a distribution which can be regarded as a random distribution, the following procedure and analysis was adopted.

The machine was to generate a distribution of 1000 'right-left' trials under the same setting (micrometer, temperature) and these trials were to be generated on two successive days in two batches of continuous recordings of similar length (Appendix 7, pp. 355-359).

Mr. Hasofer suggested the following tests:

1. Comparing the frequency distribution of 'right-left' trials in intervals of 10 trials with the frequency distribution which would be expected on the basis of the binominal distribution.
2. Comparing the strings (numbers of consecutive) 'right' trials and 'left' trials with a theoretical distribution. This is known as the 'gap analysis'.
3. By using chi square in both cases the probabilities that the empirical distribution differed from the theoretical distribution were to be calculated.

The binominal comparison resulted in a chi square of 2.87 (df = 5) with $p > 0.70$. (Appendix 7, pp. 361-362).

The gap comparison resulted in a chi square of 4.39 (df = 5) with $p \approx 0.5$ (Appendix 7, pp. 359-360).

It can be concluded from these results that the machine generated a distribution which can be regarded as a random distribution.

23. Mr. Hasofer, lecturer in mathematics, and Professor Pitman, Head of the Mathematics department, University of Tasmania, who was consulted by Mr. Hasofer, agreed that the above conclusion is justified.

TABLE 6

The 'right-left' distribution of 10 consecutive intervals of 100 trials each from the data used in the 'random' analysis

1 - 100	56	44
101 - 200	47	53
201 - 300	69	31
301 - 400	59	41
401 - 500	56	44
501 - 600	60	40
601 - 700	39	61
701 - 800	56	44
801 - 900	61	39
901 - 1000	55	45

The 'right-left' distribution in intervals of 100 (Table 6) did not indicate any systematic drift and suggested that this distribution can be expected to remain fairly stable over periods of hours and possibly days.

The data for the random analyses were obtained some time before the experimental sessions started and representing, as far as can be judged, a typical sample of the machine recording. That is with respect to technical adjustments this sample was in no substantial way different to later recordings.

It was found however that over periods of days the micrometer had to be readjusted slightly in order to remain close to a 50/50 ^{only} distribution. A systematic drift towards one direction was not noticed.

During the period when the 100 trials for random analysis were recorded, the experimenter attempted to remain unconcerned about the performance of the machine and the records were not inspected until after the 1000 trials had been completed. This of course does not rule out the possibility that the distribution was, in fact, influenced (by PK) through the experimenter, but this possibility has no immediate practical consequences because the distribution can be regarded as a random distribution.

Recording errors

As mentioned previously (p. 174), the recording unit had been tried out prior to the construction of the PK apparatus and when the inputs (to be recorded) were provided through impulses generated through the closing of contacts by a constant speed motor, the unit had worked with a high amount of reliability. For instance during a continuous run without adjustments, the unit recorded 1200 responses correctly.

When the unit was attached to the PK apparatus the following type of errors occurred with varying frequencies.

- (a) The 'one' key was not pressed or not sufficiently pressed into the adding machine.
- (b) Similarly the 'zero' key was sometimes not pressed sufficiently. However as far as is known this error

occurred in a sequence of two zeros only once, i.e. when 100 was required, occasionally only 10 was pressed in but not just '1'.

- (c) Through vibration the 'one' or the 'zero' keys pressed in these numbers more than once during one activation of these keys.
- (d) The 'add' key was not pressed or not sufficiently pressed to print the pressed in numbers.
- (e) The 'add' key was pressed longer than necessary to print one row. Therefore additional empty rows appeared on the record.
- (f) Similarly the 'sum' key was pressed longer than necessary to print one row.
- (g) The particular 'Addo X' machine printed on a few occasions the number '9' when the 'zero' key was pressed. This seemed to be an internal fault of the machine which was perhaps brought about by the vibrations during the automatic operations.

It is likely that the errors occurred because the PK apparatus provided slightly different input impulses but this is not certain because some errors may have been due to the deteriorations in the mechanical movements of the relays in the recording unit.

Fortunately the results could be interpreted without ambiguity in spite of most of the errors listed above.

Type (e) and (f) errors are irrelevant unless the spacing is used to determine where a type (a) error occurred.

Type (c) errors are irrelevant if in one set of 10 numbers 10 rows of numbers are actually printed.

One type (d) error is irrelevant unless a type (c) error occurred also. A type (d) error will result in the printing of two numbers in a row which should have been printed in a column. But without a type (c) error such a row can be interpreted correctly and without ambiguity. That is 1001 should have been printed as 100₁

Records with type (b) errors can be interpreted correctly and without ambiguity because it was not noticed at any time that both zeros in a 100 score were left out. However this is only subjective observation - nevertheless based on a large number of trials but it is impossible to say with absolute certainty that such double zero errors have not occurred.

The main error from a practical point of view was the type (a) error. To avoid this error the activation times of the solenoids concerned were slightly increased. This resulted in more vibration and slightly untidy recordings which could nevertheless be interpreted without ambiguity.

TABLE 7

Example of a recorded interval of 10 trials with various recording errors which can be corrected without ambiguity.

$$\begin{array}{r}
 1 \\
 1 \\
 1100 \\
 11 \\
 \\
 11 \\
 1000 \\
 111 \\
 \\
 11 \\
 1 \\
 \hline
 1 \\
 2248 = 208
 \end{array}$$

By changing the solenoids circuits from AC to DC, errors due to vibration were later considerably reduced.

However not all errors were eliminated but the records from the experimental analysis do not show any ambiguity.

There exists still the possibility that some errors which changed the scores could have remained unnoticed. For instance if in one interval of 10 recordings, no type (a) error occurred but if a double type (b) error occurred, then the recorded result is in disagreement with the events to be recorded. If such errors occurred they must have been exceedingly rare. Moreover such errors should not interfere too strongly with any interpretations of the results since they should occur with similar frequencies during these trials when a subject presumably tried to exert PK and during those trials when the machine ran on its own.

The PK apparatus experiment

The apparatus was used as previously described but in order to improve the disguise of the PK tests and in order to make the disguise look more reasonable to subjects who might have been interested in the technical aspects of the set up, two miniature electromagnetic coils were added and placed on both sides of the wedge (see Fig. 10 ; the coils are not included in the photograph).

Wires were connected to these coils but they were not connected to any of the operating circuits of the machine.

According to the pre-experimental tests it could be assumed that the machine generated a random distribution. There was also no sign of a systematic drift, i.e. to one side only in the 'right-left' distribution of the ball, over a considerable period (Table 6).

The following test method was made on the basis of the test procedure. Each test session was to consist of 50 trials for the subject. These will be referred to as 'subject-trials'. Before and after these 50 subject-trials the machine was to record two further sets of 50 trials. A total test session, therefore, consisted of 150 continuous trials and the subject's PK output was directed towards the middle 50 trials. PK was to be estimated by comparing the 50 subject-trials with the rest of the trials.

Although only 50 trials prior to the subject-trials were to be taken into account for the analysis, the machine was to be kept running more or less continuously and the first 50 trials of a complete test session would not normally consist of the first 50 trials after the machine had been started on a particular day. By avoiding a definite number of trials between starting the machine and the start of the test session it was hoped to avoid any warm-up effects which the machine might have.

As it was necessary to stop and start the apparatus between test sessions for such maintenance as cleaning contacts, applying oil to mechanical bearings etc., it was not possible to base the length of a fore-period (i.e. a number of trials from start of machine to start of test session) on a suitable random distribution. However it was avoided starting the machine just prior to the beginning of one test session and the machine was left running most of the time during days when test sessions took place. Two to three test sessions were conducted on several occasions while the machine continued running without interruption or adjustment.

During the 50 subject-trials the subject was not actually asked to influence the distribution of the machine. The subject was nevertheless made aware that he could influence the distribution.

The first experiment was to consist of 25 test sessions, i.e. of 3750 trials with 1250 subject-trials and 2500 machine trials.

The purpose of the experiment was to see whether any evidence for PK could be found under these conditions.

It was also hoped that the particular set-up might provide conditions which are suitable for a more stable PK output and that repetitions of PK results might be more easily possible with experimenters who do not claim to have any special abilities in parapsychological test situations.

Subjects were psychology I students who arranged testing times individually with the experimenter. They arrived at times when the machine had been running for various periods, often for hours.

Prior to the subject-trials the subject was settled in a neutral room (neither experimenter's nor subject's room) and was asked to read carefully through the following introduction to the experiments:

Measurement of Emotional Changes

Introductions

You have probably heard of Lie Detectors. The principle of such an apparatus is that bodily changes which cannot easily be controlled are measured. Bodily changes which can be used for such measurements are the heart rate, breathing rate, the skin resistance and others.

If instruments are sufficiently accurate, even a neutral question (e.g. What is your name?) and the answer to it will produce a measurable change in your skin resistance. If the question is significant (e.g. has to be answered

by a lie) the change measured is likely to be larger.

You will have a demonstration of the technique later in the year.

The Experiment:

In this experiment use is made of the fact that even comparatively neutral questions or stimuli produce a measurable change. You will be shown a series of slides and your emotional reactions to them will be measured by recording the changes in your skin resistance.

Your reactions to the slides will be compared with your reactions to an electric shock. The electric shock will be of the same strength throughout the experiment and will serve as a standard stimulus.

The comparison between shock and slide reactions is made by the use of a sensitive apparatus where an accurately ground steel ball will enter one of two lanes in a certain proportion (e.g. 40 to 60). Depending on which lane is entered, you will receive a shock or a slide will be projected. Now the change in your skin resistance will change the strength of two magnetic fields which are produced near the dividing point into two lanes. Hence a very small change in your skin resistance will change the proportion of shocks to slides. (The apparatus will be shown to you). The balance is arranged such that desirable and interesting

slides will rule out some of the shocks. However, only a sufficiently strong emotional reaction will change the slide-shock proportion in a desirable way (from your point of view, i.e. less shocks).

You must expect to receive some shocks throughout the experiments. If it should appear to you that the number of shocks rather than the number of slides increases, do not feel alarmed. This is a normal reaction for many subjects.

Instructions:

Keep the electrodes on your fingers.

Relax and watch the slides carefully.

A short light signal will be given shortly before the stimuli (shock or slide) are presented.

Remember the slide - shock proportion will depend on your emotional reactions.

The subject was left for at least five minutes undisturbed to read through this introduction. Afterwards the experimenter answered any questions which might have arisen, in the sense of the introduction, i.e. without referring to PK.

The apparatus in the experimenter's room was then shown briefly to the subject. The electromagnetic coils were pointed out. The whole introduction was conducted in a matter of fact

tons and most subjects appeared impressed with the set-up. There were no awkward questions. The subject was then conducted into the subjects' room and the shock unit was introduced. A mild shock was tried out as described for the GESP experiment (p.155). The subject was finally settled about five to 10 trials before the subject-trials were due to start.

By suitable switches the shock unit and the slide projector could be switched on and off in the subject's room. Both these stimuli were disconnected to the PK apparatus while the subject was settling in his room and while the machine trials continued. The subject was finally informed that the experiment would start in about 2 minutes and that the experimenter would return briefly prior to the first trial. As soon as the last of a set of 10 machine trials was completed the experimenter went into the subject's room and connected the two stimuli and announced that the first trial was about to begin. The door of the subject's room was then closed and the subject was left for 50 trials. In Fig. 17 a ground plan of the relevant room is shown.

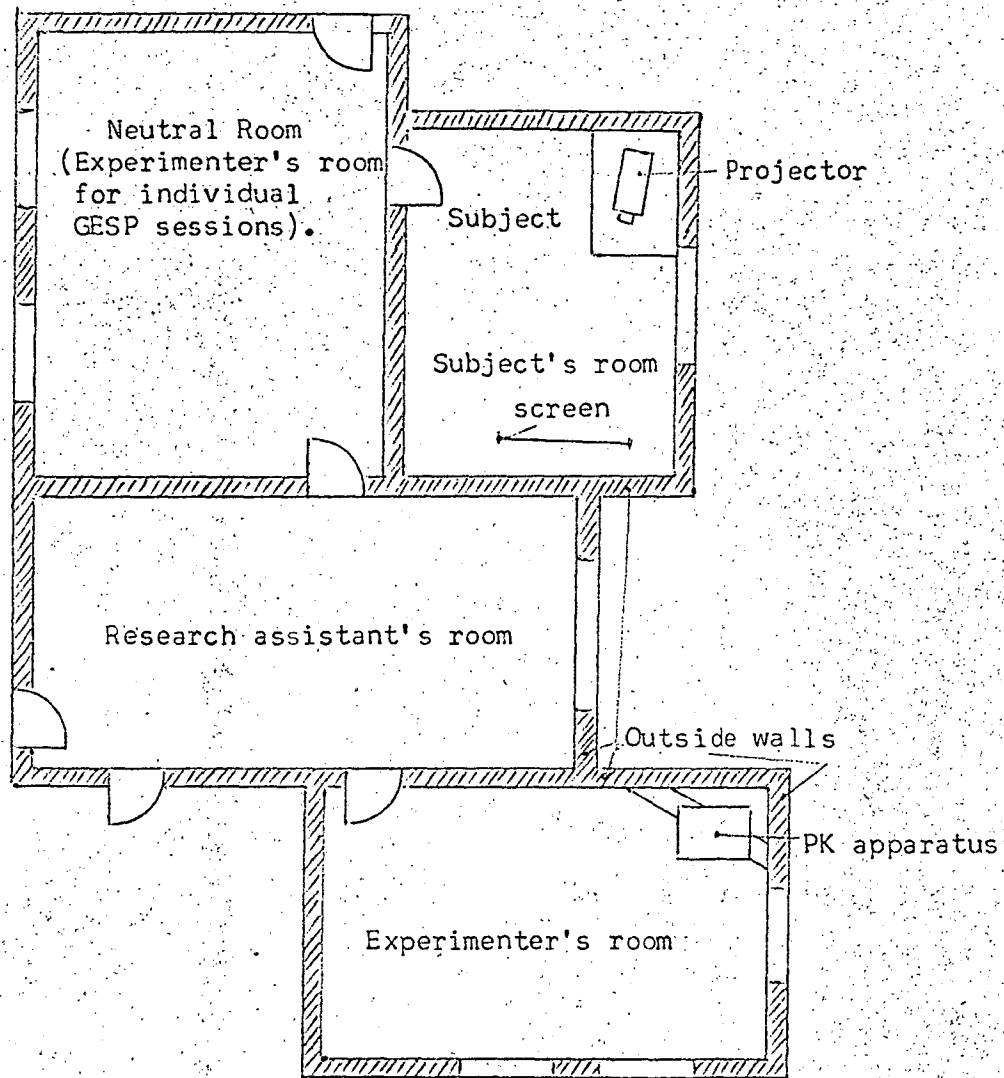


FIG. 17

Sketch of relevant rooms for the Tasmanian PK apparatus experiment. Not to scale. Shortest distance between subject and PK apparatus 5 yards approximately. During test sessions in the individual GESP experiment the experimenter was situated in the neutral room.

Minor variations which were later introduced into the procedure as additional controls will be discussed later (p.231).

Although it was expected that most of the 25 test sessions would be carried out with different students, anyone willing to volunteer a second time was to be included.

From pre-experimental trials it appeared that the right-left distribution would remain sufficiently stable to conduct most and preferably all test sessions under fairly similar right-left probabilities (e.g. these probabilities might change from 0.4 and 0.6 for right and left respectively, to 0.7 and 0.3 but they should, as far as possible, not change from say 0.1 and 0.9 to 1 and 0).

It seemed therefore possible to provide test conditions which would permit the analysis of the total experiment. The PK hypothesis was to be tested by comparing the total right-left machine distribution with the total right-left subject distribution and by applying a chi square test. The experimenter expected that the 'right' trials which resulted in a shock would decrease in the subject-trials compared with the 'right' trials in the machine trials. However a one tail test was not specified. The PK hypothesis was to be tested on this basis only but other features of interest in the results were to be investigated and analysed as far as possible.

Recording errors and other mistakes and difficulties during experimental sessions.

One subject was left by mistake for 60 subject-trials instead of 50. The 60 trials of this subject were included for analysis.

In another case the paper spool of the adding machine (part of the automatic recording unit) came to an end during the subject-trials. The spool was changed and about 7 or 8 trials were not recorded. The complete set of 10 trials which included the 7 or 8 missing trials was left out and the recording was restarted with the beginning of a new set. In this case another subject also sat through 60 trials instead of 50, but 50 trials were recorded only. During this spool exchange the PK apparatus was not interrupted.

On one and only one occasion a marked change occurred in the distribution during the subject-trials which continued in the later machine trials and after the test sessions and which had, as far as can be judged, nothing or little to do with PK. The distribution which was in this case, somewhat one sided anyway, drifted rather rapidly towards the favoured side and stayed there. In fact the ball went 109 out of 110 times to one side only. It was decided then to exclude this session from analysis and one further test session was conducted. (See also p. 218)

Although some of the printing errors described earlier (pp. 204-206) appeared through the recording of the test sessions they did not introduce any ambiguity.

Although it can be assumed on the basis of the randomness tests (p. 201) that the machine generated 'by itself' a random distribution at all times, it is difficult to test the distributions of the test sessions for randomness. This is due to the fact that one test session does not include a sufficient number of trials, yet when test sessions are combined, the tests for randomness are only meaningful if it can be assumed that the probabilities for right and left trials remain unchanged during the test sessions. The analysis of 1000 trials suggested that the probabilities will not change rapidly but it cannot be assumed that the probabilities for right and left trials remain unchanged during the test sessions. The analysis of 1000 trials suggested that the probabilities will not change rapidly but it cannot be assumed that they stayed the same over days and weeks.

Indeed some slow changes had to be corrected by altering the micrometer setting. But such alterations were only carried out between test sessions.

Results:

In the total experiment 26 test sessions were carried out. One test session was excluded and the remaining test sessions consisted of 2500 machine-trials and 1270 subject-trials but only 1260 subject-trials were recorded.

Twenty-three psychology I students participated as subjects and two of these participated twice.

The total right-left distribution of the machine-trials was 1597 and 903 respectively and 730 and 530 for subject-trials. Testing the difference between subject-trials and machine-trials by chi square (Appendix 8, p. 363), it was found that the null hypothesis can be rejected with $p < 0.001$ (chi square ≈ 12 ; $df = 1$).

Since no rapid change in the distribution (as occurred in the one case excluded) was anticipated, no rules for exclusion of results were established prior to the experiment. However if the difference is tested in a similar way as above for all 26 test sessions, i.e. with 150 trials (previously omitted) included, chi square is still = 11.33, i.e. the null hypothesis can still be rejected with $p < 0.001$ (Appendix 8, p. 363).

Since the inclusion does not affect the PK hypothesis in any substantial way this test session has been left out in further analysis as it is likely to distort any secondary evidence.

It seems that the PK hypothesis can be accepted. Various counter hypotheses and further improvements in controls will be discussed but these considerations do not lead to a satisfactory explanation of the results by non-parapsychological means.

To check whether any general drift occurred in the machine distribution throughout the complete experiment, the total machine trials recorded before the subject-trials were compared with the total machine-trials recorded after the subject trials. The right-left distribution for the 'fore' trials was 800 and 450 respectively, and 797 and 453 for the 'post' trials. Testing the difference by chi square (Appendix 8, p. 364), the probability that the

two machine-trial distributions are from the same population, is 0.9 (chi square = 0.016 ; $df = 1$).

Quite apart from the probability calculation the change in the distribution during subject-trials is quite apparent in Table 8.

TABLE 8

Summary of right-left distributions in 'fore' and 'post' machine-trials and in subject-trials.

	Shock	Slide	
	Right	Left	Total
Machine fore trials	600	450	1250
Subject-trials	730	530	1260
Machine post trials	797	453	1250
Total	2327	1433	3760

The chi square analysis to test the difference between the machine-trials and the subject-trials with respect to the right-left distributions was also carried out for each test session (Appendix 8, pp. 365-371).

The chi square values are listed in Table 9. This Table also includes information about the QESP results obtained by those ²⁴PK volunteers who participated in other tests.

24. Names of subjects are recorded in the original test material.

TABLE 9

Summary of parapsychological results of subjects who participated
in the Tasmanian PK - machine tests.

Sub- ject	date 1961	PK machine tests			ESP group tests		ESP individ. tests	
		chi square between S. trials & machine trials	P	+ = more slides - = more shocks 0 = no change	Direct hit dev- iations in 4 runs	-1)dis- place- ment in 4 runs	Direct hit dev- iations in 4 runs	-1)dis- placement in 4 runs
A	4.5.	2.67	<0.2	+				
B	5.5	0.12		+				
C	11.5.	0		0	+3	+4	-2	+2
D	12.5.	0		0	+3	0	+6	+4
E	16.6.	0.53		+	+6	0	+2	0
F	22.6.	0.13		-	+1	-5	+4	+1
G	22.6.	1.33		-	+6	-6	-3	+10
H	23.6.	0.13		+	-1	-4	+5	-4
I	28.6.	4.32	<0.05	+	-5	-2	+1	+5
J	28.6.	5.88	<0.02	+	+3	-1	+1	+4
K	6.7.	1.95	<0.2	-	+1	+4	+1	-2
L	6.7.	3.02	<0.1	+	-1	+3		
M	6.7.	2.52	<0.2	+	+3	+2		
N	6.7.	5.66	<0.02	+				
O	7.7	4.11	<0.05	+				
O	3.8	4.32	<0.05	-				
P	7.7	5.87	<0.02	+				
P	11.8.	3.57	<0.1	-				

TABLE 9 (cont.)

Sub- ject	date	PK machine tests			GSP group tests		GSP individ. tests	
		chi square between S- trials & machine trials	P	+ = more slides - = more checks 0 = no change	direct hit dev- iations in 4 runs	(-1)dis- place- ment in 4 runs	direct hit dev- iations in 4 runs	(-1)dis- placement in 4 runs
Q	12.7.	0.12		•	+3	-2	+3	+1
R	12.7.	0.12		+	+3	+4	0	+3
S	13.7.	2.23	0.2	+	+3	-2		
T	13.7.	0.57		+	+1	+12	+3	+7
U	20.7.	4.82	0.03	+				
V	4.9.	0.004		+	+1	-1		
W	11.8.	0.62		-	+5	-3	+5	-2

Total 54.67

The two subjects who participated twice did not carry out their two tests in succession but volunteered on different days.

The chi square total = 54.67 (df = 25) indicates a significant deviation with $p < 0.001$. To obtain more information about the amount of success at various stages of the test session the total right-left distribution for the first, second, ... fifth set of 10 subject-trials was established (Table 10).

These totals were compared with the total right-left machine-trials and chi square was calculated for each set (Appendix B, pp. 372-373).

TABLE 10

Total right-left distributions of all first, second, ... fifth sets of 10 subject-trials per test session.

Subject-trials	Shock right	Slide left a_2	Total				chi square between 5 trials & machine trials	P	$a_2 - \frac{b}{10}$	$a_2 - \frac{a_1+a_3}{2}$
1-10	151	99	250				1.21	<0.3	+8.7	+13.5
11-20	141	109	250	Total machine-trials			5.46	<0.02	+18.7	+20.5
21-30	138	112	250	Right	Left	Total	7.56	<0.01	+21.7	+14
31-40	151	99	250	1597	903	2500	1.21	<0.3	+8.7	+10.5
41-50	149	103	250				3.21	<0.1	+14.7	+14

a_1 = totals of first, second, ... fifth left machine 'fore' trials.

a_3 = totals of first, second, ... fifth left machine 'post' trials.

The differences between the number of left-trials expected on the basis of the total number of left machine-trials ($b/10$) and the actual number of left subject-trials (a_2) were also included in Table 10 and represented graphically in Fig. 18.

Finally the differences between the number of left-trials expected on the basis of the corresponding two sets of 'fore' and 'post' left machine-trials ($a_1/2 + a_3/2$) and the actual number of left subject-trials (a_2) were also included.

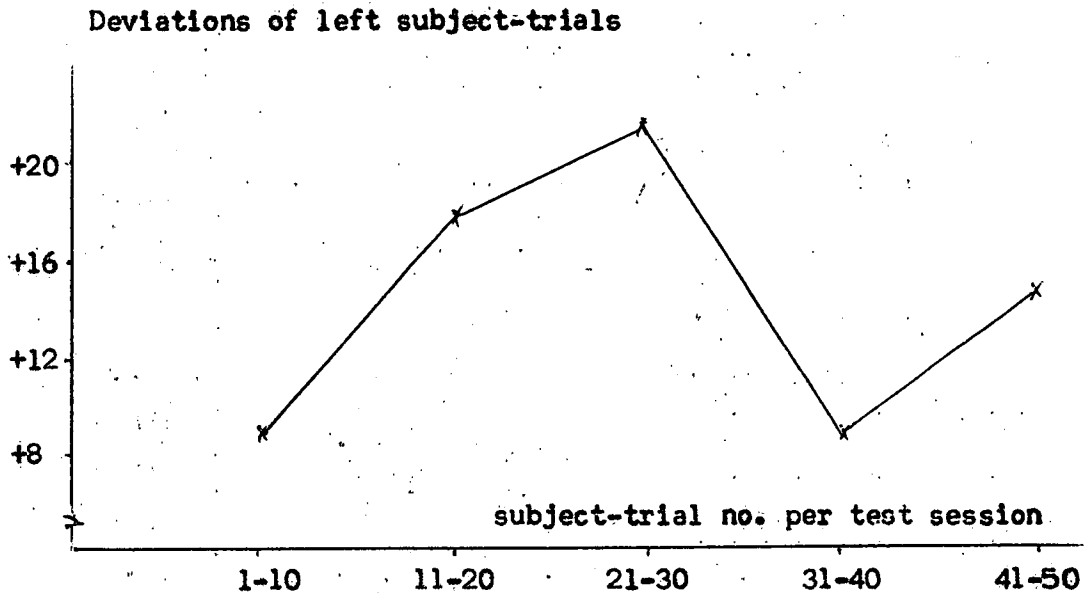


FIG. 18

Graph of left subject-trial deviations per interval of 10 trials in each test session.

23 subjects participated in 25 test sessions consisting of 50 subject-trials each. The deviations are calculated by subtracting the expected number of left trials based on the total number of machine trials from each interval total of left subject-trials.

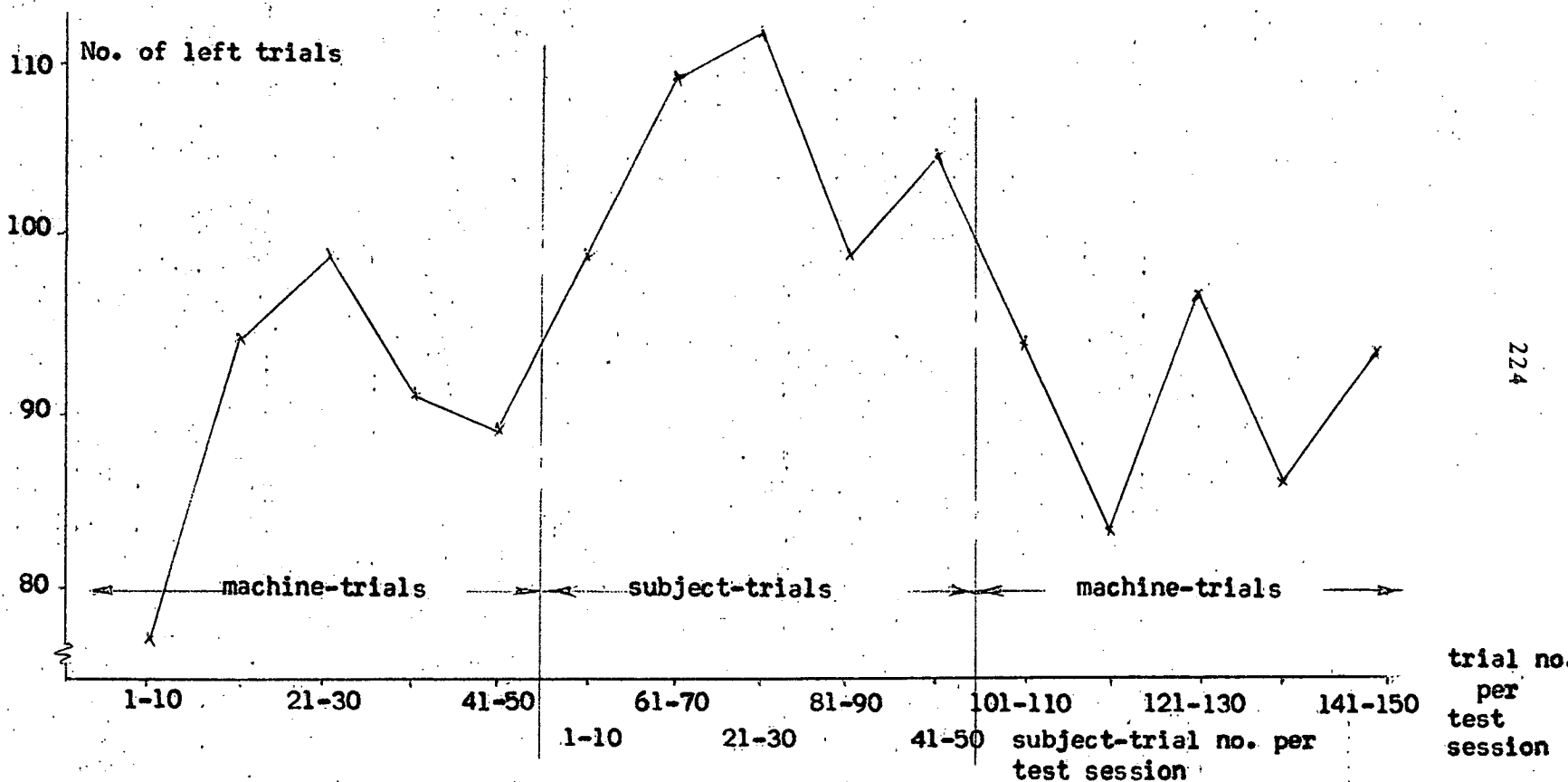


FIG. 19

Graph of left trials per interval of 10 trials in each test session.

23 subjects participated in 25 test sessions, each consisting of 100 machine-trials and 50 subject-trials.

It is difficult to say which set of deviations is more representative of the amount of PK at various stages of the test sessions because the machine may have favoured some deviations and disfavoured correspondingly some others.

Taking the total machine trials (Fig. 18) as the basis for comparison probably provides a better approximation than when - in a sense - arbitrary sets of trials are compared, i.e. in the 150 trials of each test session trials 1-10 (machine-trials) are added to trials 101-110 (machine-trials) and the mean of these sets is compared with trials 51-60 (subject-trials). There is no compelling reason why trials 51-60 should not be compared with trials 11-20 and trials 111-120. Nevertheless it is interesting to note that in both kinds of deviations calculated, all values are positive.

Fig. 19 provides a graphical summary of the total experiment including the 'fore' and 'post' machine-trials. The graph is drawn on the basis of left-trials only. During the left-trials the subjects did not receive shocks but saw slides instead. The equivalent graph for the right-trials could be obtained by rotating the graph of left-trials around a horizontal axis by 180° .

The fifteen results of total left trials per interval of 10 trials were separated into ten results (machine-trials) making up group 1, and five results (subject-trials) making up group 2. Group 1 and group 2 were compared by analysis of variance and it was

found that the null hypothesis can be rejected on the basis of this analysis close to the 0.001 level ($F = 16.79$; $p = 0.01$ for $F = 9.07$ and $p = 0.001$ for $F = 17.61$). It is not certain whether all assumptions for this analysis can be regarded as fulfilled but on the basis of the random analysis and on inspection of the experimental results it can be assumed that the necessary assumptions are approximately fulfilled and this should be sufficient (McNemar, 1955, p.255) for the purpose (p. 227), of this additional analysis.

Discussions

Considering Table 8 and the more detailed Table 9 the rejection of the null hypothesis on the basis of $p < 0.001$ (Appendix 3, p. 363) seems justified. A decline as in the GESP group test (Fig. 3, p. 140) did not occur (Fig. 18) but it seems doubtful whether a similar decline should be expected under the conditions of this experiment. There are at least two factors which could have delayed or counteracted the decline of the PK results.

1. The disguise could have delayed the onset or correct direction of PK. It may be noted that it appears at least as if in the disguised GESP test, parapsychological processes may have come into the test situation more readily than in the PK experiment. On the basis of this assumption a delay in the maximum PK output could be expected.

2. The reinforcing stimuli may have interfered with the decline in the PK output. There is at least a superficial similarity between the unreinforced GESP test and this PK test with respect to the results at various stages of each test session (Fig. 6 and Fig. 18).

Experiments with the PK apparatus which were carried out at the Parapsychology Laboratory, Duke University, under less satisfactory controls and under somewhat different test conditions will be discussed separately (p. 244). It is doubtful whether the Duke results can be regarded as supporting evidence but it seems correct to say that they are not in conflict with Tasmanian results.

Although the analysis by chi square was specified prior to the experiment it may appear on inspecting Fig. 19 that the similarity between the total 'fore' and 'post' machine-trials is perhaps by chance too close. The change between the first interval and the second interval of 10 trials per test session is larger than between any other adjacent intervals and hence it seems possible that the probability based on chi square is somewhat too favourable.

Because of these considerations an additional analysis of variance was carried out and although the calculated probability is not quite as small as in the chi square analysis it is of the same order of magnitude (Appendix B, pp. 374-376).

It seems then that the rejection of the null hypothesis can be accepted and the main concern which remains is to consider whether other counter hypotheses could account for the results.

The experiment was conducted on a one experimenter basis and it cannot be claimed that fraud would not have been impossible in principle. Nevertheless the machine set up might make it possible to develop a procedure which might satisfy critics like Hansel and G. Price but which presumably would not change the experimental situation from the subject's point of view.

The principle requirements would be to engage a second parallel recording unit such that results are recorded in duplicate and to work out some system which prevents any interference with the apparatus and test conditions (e.g. room temperature) during test sessions. At present it can at least be claimed that the automatic recording unit prevented any recording errors which might have favoured the PK hypothesis. Other possible counter hypotheses may be listed as follows:

1. The results were due to a warm up effect of the apparatus.

The temperature of the PK apparatus was thermostatically controlled in Tasmania. A thermostat which is commercially available was placed underneath the cover of the PK apparatus close to the release channel. The thermostat triggered two electric radiators which were placed as far from the PK apparatus as possible and which did not reflect into the direction of the apparatus. The

experimenter's room in which the apparatus was situated was kept at a nominal temperature of 73°F .

It was found that with the aid of a thermometer which allowed readings up to 0.1 degree F and which was placed on the apparatus, that the temperature varied by a maximum of 1° with 73° as approximate mean.

During the experimental sessions the temperature did not rise above 73.5° and it is unlikely that the outside temperature ever reached this figure at that time of the year.

The two radiators seemed sufficient to keep up the required temperature when outside temperatures were considerably below 73° . Even though the machine was switched off overnight the temperature control continued throughout that period during which pre-tests and experiments took place. It is unlikely that, under these conditions, changes occurred in the right-left distribution which depended on the number of trials from the first start of the machine. No such systematic changes were observed by the experimenters.

The attempt to reduce any systematic warm up effects of the machine, if they existed, by varying the lengths of the 'fore' periods before test sessions were started, is regarded by this writer as adequate. However it would

be desirable to change these 'fore' periods on a more formal basis by the use of a suitable random distribution. To make this practicable the sections of the machine which require maintenance will have to be improved.

It seems unlikely that any warm-up effect existed which changed the right-left distribution but if such an effect was present it must have been small and it seems even more unlikely that it could have changed the distributions recorded during test sessions in a systematic way.

2. The results were due to physical changes which occurred during the test sessions.

This hypothesis can be considered in two parts:

- (a) changes which were part of the test procedure might have caused changes in the distribution.
- (b) less specific changes which might have occurred under the experimental conditions caused changes in the distribution.

Under (a) it may be noted that according to the experimental procedure the reinforcing stimuli were switched on just prior to the subject-trials. The electric power for these stimuli was provided in the subject's room at some distance from the PK apparatus which only via relays switched the stimuli at the appropriate times and for a certain length of time.

It seems unlikely that the difference between activating a relay which closes a separate circuit which in one case was alive and in another case not, should make a difference to the 'right-left' distribution.

Also the reinforcing stimuli were left on during the 50 machine-trials following the 50 subject-trials. (Electrodes were removed from the subject).

Under (a) it may also be noted that subjects observed the machine for one or two trials about 10 to 15 trials prior to the subject-trials.

Since the apparatus was covered and rested on steel brackets bolted onto two stone walls it is unlikely that this could have changed the distribution. The adjacent room was occupied by a research assistant of the psychology department and was heated. At the time when the experiments took place there was no noticeable temperature difference between the two rooms.

Five of the subjects were also asked to have another look at the apparatus after the subject-trials. In these cases for instance, physical conditions were provided for the machine trials which were similar to those prior to the subject-trials but this did not lead to any noticeable changes.

For similar reasons it is unlikely that less specific changes (b) had an influence on the distribution. The experimenter's room

was also used for other work and it cannot be claimed that the machine was kept in a room which itself was isolated from human habitation, but the apparatus was so situated that accidental interference was unlikely and did not, as far as is known, take place at any time. The presence of the machine in the experimenter's room made it, on the other hand, also impossible for someone else to disturb intentionally or unintentionally the distribution without the experimenter's knowledge.

It may be desirable in future to house the apparatus in a special room (if available) but this isolation of the machine is only practicable if the machine can run over long periods without maintenance or readjustment.

Such alterations could not have been introduced at the time when the experiment took place but it seems to this author that the existing conditions were sufficiently adequate to rule out counter-hypotheses based on physical changes. Hence it seems justified to accept the PK hypothesis.

Although the results reached a level of significance which seems satisfactory it cannot be claimed that the test conditions were such that all subjects scored in a similar way. Table 9 indicates that probably less than half of the subjects contributed markedly towards the general change in the right-left distribution. On the other hand the change can be regarded as being based on 6 or more subjects, i.e. the total change is not

likely to be based on one or two subjects only.

As has been discussed previously the design of the experiment tended to move the experimenter further back out of the subject's experimental field, yet the possibility that the total change occurred because of the experimenter's PK quite independently of the subjects, cannot be completely ruled out. The comparatively large change from the first to the second interval of 10 trials per test session (Fig. 19) could be attributed to the experimenter's desire to avoid a too one-sided distribution of the right-left trials. Also in this experimental situation the experimenter-subject relationship may still have played a role. Nevertheless the disguise was accepted without hesitation by the students and at least the open introduction of a controversial topic : parapsychology, was avoided.

It may be noted that the total distribution of the right-left machine-trials did not come out close to 50/50 but was, on an average, 64/36. That is, on the basis of this distribution subjects would receive more shocks than slides and although this distribution changed from 64/36 to 42/58 when subjects were present (during subject-trials), the number of shocks presented was still greater than the amount of slides.

The somewhat one-sided distribution of machine-trials was unintentional. It seemed that the apparatus could be more easily adjusted to generate the 64/36 distribution than a 50/50 distribution. However after a number of test sessions had

been conducted with the one sided distribution of machine-trials, no attempts were made to counterbalance the previous tests by a new micrometer setting likely to result in a right-left distribution of opposite bias.

The one-sidedness of the right-left distribution has no bearing on the statistical analysis presented here which does not require a 50/50 distribution in the first place. However the greater number of shocks compared to slides may have been responsible for a particular psychological test situation which had perhaps beneficial effects in the PK tasks.

Generally it would be expected that PK can operate more easily if a 50/50 distribution is to be changed rather than a 64/36 distribution. Nevertheless the comparatively large number of shocks which were given during the early subject trials may have induced subjects to change the distribution at a greater rate than if the distribution had been biased in the opposite direction.

Suggestions for changes of the apparatus have so far dealt with details which might improve the reliability of the various operations involved, including the recording operations. It should perhaps also be discussed whether more fundamental changes should be considered. In the present set up the speed with which trials can be conducted is governed by the cycle frequency of the alternating current used to operate

the control unit. Small variations are possible (e.g. from 50 c/s to 60 c/s) but it is unlikely that the number of trials could be doubled without running into a number of difficulties.

A question which is relevant and which should be asked is whether it is necessary to use one and the same ball. The answer is that it is probably not necessary, but desirable for continuous unattended operation of the machine. From all pre-experimental and experimental trials conducted with the apparatus it would appear to this author that the maximum unattended running time was, under favourable conditions, about 24 hours. Whether 'continuous' and unattended trials under such limitations warrant the comparatively slow and complicated return mechanism of one ball is difficult to say. Nevertheless the aim to have continuous background recording, possibly of some considerable length, against which any PK effects may be studied, seems sound enough particularly if PK in wrong direction (p. 177) is to be considered.

The speed at which an experiment proceeds can have an effect on the kind of parapsychological events which are obtained and the amount of displacement may vary (Soal & Bateman, 1934) but so far it is difficult to say whether better PK results will be obtained if trials proceed at a faster rate than in the experiment discussed here.

It is also doubtful whether a trial with reinforcements can be compared with a trial without reinforcements. Hence it seems to this writer that it would be desirable to leave the principle operations and the time sequence of the PK apparatus unaltered, but to improve the reliability further than has been possible so far. This should enable the investigator to make full use of the desirable features of the design of the apparatus and a continuous background recording over a period of days is one of these.

Some considerations about future work with the PK apparatus.

If, as the results will suggest, the principle of reinforcement of correct responses has research possibilities, and if it is to be applied intelligently, it is necessary to consider the following questions:

Are there any stimuli which are generally accepted as positive or negative by a limited sample of the total population, say by first year psychology students?

It was not attempted in any systematic way to find such general stimuli but this question was further pursued by supplying favourite pieces of music as GESP targets. (see p. 264) which were prepared for each individual subject. Although no personality tests were introduced, this part of the research provided the closest links with the personality studies of prospective subjects for parapsychological experiments. And it

seems that here a considerable expansion of research is possible and probably even desirable from the point of view of contemporary psychology.

The problems which should be investigated are:

1. What available psychological tests point out areas of interest in individual subjects (e.g. Kuder, 1957, Layton, 1960)? What tests measure the intensity of interest, pleasure, pain or other subjective variables of positive or negative orientation towards the test stimuli.
2. What tests could be constructed to investigate the problems outlined above if the existing tests are insufficient? Schneidlor's and Humphrey's approach, reviewed by Humphrey (1951) may well be extended through tests to estimate the 'opermindedness' of subjects (Rokeach, 1960).

Electric shocks are generally considered to be undesirable by subjects at an overt level. If any psychoanalytical theories are accepted then it must be expected that some subjects will nevertheless desire this kind of punishment at a subconscious level. But even without such theories the problem of finding a stimulus which is generally experienced as negative, is far more difficult than it may appear.

Comparatively small electric shocks are not entirely without danger as was discussed by Keesenhoven (1949) and by Picken (1961). But it is also difficult, if not impossible, to persuade volunteer

students to accept very unpleasant (although physiologically safe), shocks. Yet if the intensity of the shock is of the order of 'just unpleasant' (according to the subject's estimation) it is by no means certain whether a group of such subjects will tend to avoid the response which is associated with the shock. Indeed a small shock may reinforce the response which leads to the shock (McGeach, 1942).

In the present study no real attempts were made to solve this problem. Yet observations showed that this problem certainly existed. Some subjects reported after the experiments that during the later part of the experiment they could not feel the shock when it was supposed to come on. (This was not due to technical failures). Physiological measurements such as GSR measurements did not lead to any result which obviously agreed with the overt verbal responses of the subjects, and verbal responses had to be taken as the guiding information to determine the intensity of the shock for each individual subject.

If a sufficiently large group of subjects is available initially it may be possible to select a group on the basis of conditioned GSR responses. (US = electric shock, Cs = buzzer). But - as was noticed by the author during experimental courses with students - these conditioned responses may vary considerably depending on whether an audience is present or not and on who works as experimenter.

McElroy and Brown's earlier attempts (1950) to use electric shocks in an ESP experiment led to results which have some similarity with those reported here except that in their group experiment - but not in the individual test reported - there was more evidence of decline.

It seems then that the reactions of any group of subjects to low intensity shocks may vary to a large extent.

Unfortunately the situation is not much better with respect to stimuli which may be regarded as positive or desirable.

Color slides have become the subject of some ridicule²⁵ (Wild life in suburbia) and cannot be expected to be viewed by students with interest unless particularly 'fascinating' photos are selected.

Fisk and West (1955) used sexually toned objects as stimuli with some success in an ESP experiment. The display of erotic stimuli has not been followed up in parapsychology, which may be due to the conventional outlook against the use of such material, but which may also be due to legislation in various countries which makes the display of certain material unlawful. Moreover departments where research with erotic material

25. A sound recording which is commercially available in Australia.

is carried out are likely to be criticised publicly through newspapers. Because of these considerations the present writer did not pursue the possibility of presenting sexual stimuli as reinforcements.

It seems quite clear however that a wide area of promising research is left unexplored (see Dingwall, p. 231 in Wolstenholme & Miller (Eds.), Ciba Foundation symposium, 1956), and it is urged that this question should be taken up by departments or institutions which have either previous experience with the material in question (e.g. Clark, 1932, 1933), or which are less likely to be criticised publicly (e.g. medical and psycho-analytical institutes).

The slides which were used as stimuli in the experiment consisted of a selection of 30 mainly European travel slides. It can hardly be claimed that they were uniformly perceived as desirable stimuli but it seems reasonable to expect that they prevented excessive boredom during the test. It can at least be claimed that the average psychology student was presented with a test situation which, without any stimulating introduction by the experimenter, was more interesting than the standard PK or ESP test. It may of course be argued that the interest is created (in the card or dice tests) through the experimenter but in this research plan one main feature was to eliminate or to minimize the influence of the experimenter on the research situation.

Nevertheless to increase the effectiveness of reinforcing stimuli, further individual personality differences should be taken into account and stimuli should be varied accordingly.

The testing of 'PK times' is of particular interest and although no experiments were possible in the time available, a brief discussion may be justified.

Most PK measurements are of a discrete and not of a continuous kind. Forwald's placement experiments (1954a) are the main exception and should, because of this, deserve more attention.

Measurements on a continuum have the disadvantage that disagreement between observers may occur. This can be overcome for time intervals by semi-automatic mechanisms which have, in the end, discrete and unambiguous units of sufficiently small size.

One thousandth of a second is sufficiently small if one and the same record of the time interval, i.e. recorded on tape by two signals which in turn can start and stop the counter - can produce different decatron counts.

It cannot be claimed at present that a stable system (see p. 84) can be changed by PK. Hence it seems necessary to be able to record changes through fine measurements.

In the above example of the tape recorded time interval no changes could be measured if the decatron counter operated in units of 0.1 seconds only. The same reading would be obtained each time and changes due to PK would be most unlikely (although certainly desirable in principle).

If the changes which occur from trial to trial and which are presumably not due to PK, are fairly large in terms of the smallest unit of measurement, then it is more likely that any additional changes due to PK will show up in more than one unit and over many trials a PK component may be more clearly extracted from these kinds of test measurements.

In an experiment which this author carried out under the direction of A. L. McAuley (1955)²⁶ significant results were obtained when subjects tried to influence the speed of the movement of a very small physical system known as a microbalance. Since the time measurements for the movements through a certain distance were carried out by the experimenter (operating a stop-watch) the results can be explained by GESP instead of PK. That is, the experimenter who did not know the fast slow sequence which the subject attempted to obtain, might have been unconsciously induced through GESP to operate the stop-watch in agreement with the subject's wishes.

In this setting it was observed that a particular subject who felt highly emotional about this task managed to 'produce' times which were quite outside the range of other experimental and non-experimental trials. The changes were so large that they could not be explained in terms of unconscious influences through ESP. However the times per trial did not correspond to the instruction sequence given to the subject.

26. Head of the physics department, University of Tasmania.

If these changes were due to PK and not an artifact then it ought to be possible to produce PK distributions which are different from non-PK distributions of the system involved. At the same time the PK distribution would not necessarily have to agree with the distribution of specific PK instructions (e.g. fast, slow, slow ...).

Time intervals may be particularly suitable for such investigations and this possibility influenced the design of the PK experiments described here.

The method of continuous variable trials has been discussed by McConnell (1958) and it was pointed out that under certain circumstances this scoring technique may be less sensitive than the binomial technique. This is correct for clockcard tests (Mitchell Fisk, 1954) but this situation is different from the time interval case discussed above. In the clockcard tests, ESP effects which are either too weak or too misdirected to result in a full hit, still score although of course, less than a full hit, i.e. partial success is not wasted. But in the clockcard test (and other ESP tests) it cannot be shown that ESP effect 'A' which scored a full hit, was stronger than the ESP effect 'B' which also scored a full hit but only 'just'. This sort of excess strength might show up in the 'PK-time' tests if the time intervals are measured in sufficiently small units.

The Duke PK experiments with the PK apparatus.

The PK apparatus and the automatic recording unit from Tasmania was reconstructed at Duke and put into operation under considerable difficulties. Although the correct voltage could eventually be supplied, the available current was reduced in Ampere and some of the solenoids operating the recording unit never worked completely satisfactorily. There was also no opportunity to introduce temperature controls or to house the apparatus on a special room. It was situated in the experimenter's room (203C in the Parapsychology Laboratory, Duke University) and rested on a skirting board of a wall. It was hoped that this arrangement eliminated or reduced small movements of the apparatus due to walking or standing on the floor of the same room, but such movements cannot be ruled out altogether. More obvious drifts in the right-left distribution and less ambiguous recording errors occurred and could not be eliminated under existing circumstances. Nevertheless at times distribution seemed to remain reasonably stable for one or two hours and it was decided to test 10 subjects.

There was no suitable automatic slide projector or shock unit to provide reinforcing stimuli. Also the necessary interval timers were not available.

It was also impracticable to disguise the parapsychological nature of the test in the Duke laboratory.

The following test procedure was adopted:

After 30 machine-trials 60 subject-trials were to be recorded which were again to be followed by 30 machine-trials within one test session. The 60 subject-trials were to consist of two halves of 30 trials during which the subject wished the ball to go in opposite directions (right, left). The direction to be shown first was to alternate from test session to test session.

The internal control within the subject-trials should be sufficient to exclude other explanations of significant changes but the machine-trials were included in the test session to enable the experimenter to decide whether a test session was sufficiently stable to be accepted.

The following rules were established prior to the experimental sessions. A test session was to be excluded under two conditions, each sufficient on its own.

1. The 30 'fore' or 'post' machine-trials only included 3 or less trials of one side. If this happened in the 'fore' trials the subject-trials were not to be carried out unless there was sufficient time to readjust the micrometer and to run another 30 machine-trials before the subject-trials etc. If the 30 'post' machine-trials^{session} reached a one sided distribution of 3 or less the whole test was to be excluded.
2. If in a whole test session (including machine-trials) more than 5 ambiguous errors occurred, this session was to be

excluded. Ambiguous refers to errors as printed (or not printed) on the paper roll of the recording unit, i.e. a trial not printed is an ambiguous error in the record, but some other form of error in the records can be corrected without ambiguity and these were not counted here (pp. 204-207). Since the experimenter was present during the total test sessions, ambiguous errors were corrected by the experimenter on the running record sheet by, for instance, writing down the appropriate number not printed. But if five or more corrections had been made the complete test session was to be excluded.

Within the available time only six test sessions were successfully (from a technical point of view) completed. Another five test sessions were started and discontinued at various stages of the experiment due to extreme changes in the distribution, or recording breakdowns.

Obviously the test conditions were not quite satisfactory and any evidence for PK must be viewed on a somewhat speculative basis. Nevertheless as a background to the Tasmanian PK experiments it seems desirable to include the results of the six test sessions here.

Because of breakdowns on some intermittent test the six test sessions did not start with alternate target directions and four subjects started with the 'right' target and only two with the 'left'

target (Appendix 9, p.377).

The 360 subject-trials were distributed as follows:

TABLE 11

'Right-left' distribution of 360 subject-trials during right
and left targets.

	Right	Left
target right	83	97
target left	65	115

Testing the difference by chi square $p < 0.055$ (chi square = 3.7 ; $df = 1$). On the basis of the unsatisfactory test controls it would be unjustified to consider this result as suggestive evidence for PK. But it seems justified to note that this result is not in disagreement with the experimental evidence for PK from the Tasmanian test sessions.

A standard PK test with dice

Throughout the author's stay at the Parapsychology Laboratory, Duke University, PK tests were conducted with one of the motor driven cages (Rhine & Pratt, 1957, p.165). These tests were mainly conducted to meet volunteer students in a simple experimental situation. Some staff members of the laboratory also participated on a few occasions. Because of the introductory nature of these tests the stopping point was not fixed until about half of the tests had been completed.

Subjects aimed for specific die faces and the same number of trials was carried out for each face.

Mainly because one subject was interesting in carrying out low aim tests, a smaller number of equal trials was carried out for each target face which, under low aim conditions, was to be avoided.

The experimenter was also interested to observe the subject's reaction to continuous information about their positive or negative deviations during the test session. One test session consisted of 30 runs of 24 trials each. Initially subjects were also given the choice of selecting a target number if they had any preferences but once the stopping point was fixed and once the necessary trials for each target ^{face} were fixed, this choice became more and more restricted.

Throughout the experiment the motor driven cages were used with six dice (thrown simultaneously) of equal size. Subjects

were informed about their deviations by the following method.

If, for example, the number 5 was the target face in a high aim test, one die with this face up was to be expected in one throw by chance. Hence, if in the first throw this target appeared three times, the number +2 was announced to the subject and entered on the scoring sheet. If at the next throw the target face did not appear at all, the number +1 was announced to the subject, where plus one is now the deviation after two trials.

In low aim tests this procedure was not changed but it was made clear to subjects that a positive deviation of say +3 which could occur early in the test session was the sort of thing they did not want and that they wanted to end up with a deviation of say -30. The length of one test session was indicated to subjects on the scoring sheet.

Subjects were encouraged to check the scoring of the experimenter but some preferred to sit back and did not observe the die faces, nevertheless the scoring sheet was shown to all subjects during the test in order to indicate to them how far the test session had progressed.

The scoring method was demonstrated to subjects participating for the first time. The experimenter was interested to observe subjects under increasing positive and negative deviations. This interest is based on the GESP (individual tests) and on the Tasmanian PK test conditions in which subjects obtained similar though less precise information about their progress in the tests.

Results:

In the high aim section 900 runs (150 runs per die face) were carried out and in the low aim section 180 runs (30 runs per die face).

For the high aim section a positive deviation of +135 occurred. For the low aim section this deviation was +3 which is negative in the sense of the low aim. The total deviation amounted therefore to +132. The probability that this deviation occurred by chance is less than 0.03 (CR = 2.2). The total deviations (Appendix 10, pp. 378-380) for five stages of each test session are shown in Fig. 20.

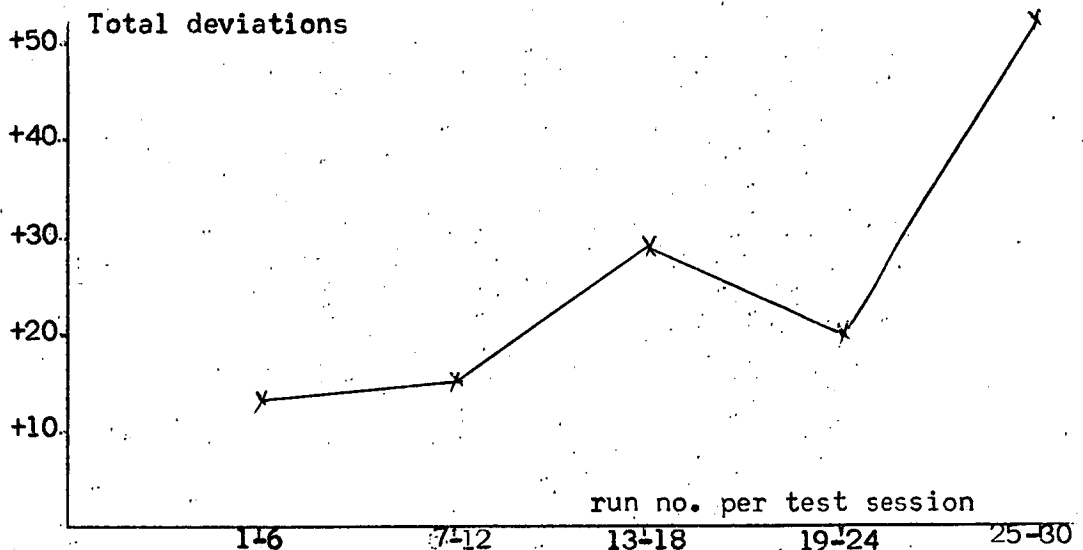


FIG. 20 Graph of total interval deviations in 36 test sessions with the motor driven cage for dice.

Each interval consists of 6 PK runs per test session. Total deviation in 1080 runs = +132.

One subject who participated in 11 test sessions with 330 runs (however not equally distributed over all target faces, although approximately so) obtained a negative deviation of 63 (CR = 1.94; $p = 0.055$). The absolute total of all eleven deviations from the eleven test sessions of this subject reached 151 (Appendix 10, pp. 378-380).

Discussions:

The total result may be called suggestive or marginally significant but the selection of a stopping point half way through the experiment and the possibility of scoring errors make the PK hypothesis doubtful.

The negative deviation of 63 for one subject alone over 330 runs suggests nevertheless that PK processes changed the distribution of the die faces.

From the results however, it seems difficult to extract further evidence for PK in a way that can be accepted as methodologically sound.

On the basis of subjective observations there was no sign that subjects felt particularly distressed if their deviations accumulated in the wrong direction. It is possible though, that subjects tended to reach higher deviations (for both positive and negative deviations) than they would be expected to reach in situations where this information is not provided. But this is at best, a reasonable speculation.

It is interesting to note (Fig. 20) that the last of the intervals, each consisting of 144 trials, i.e. of 24 throws of 6 dice, has the highest deviation. While it is only possible on a speculative basis to relate this to the particular test conditions (calling out progressive deviations), it seems desirable to try out these test conditions again by incorporating them into other future experiments.

Physiological measurements and parapsychological processes.

An average score of 6 in an ESP experiment may support the ESP hypothesis; at any rate given a sufficient number of runs, the null hypothesis may be rejected on the basis of a very small probability.

Yet among six hits recorded on an average, only one can be attributed to ESP. Obviously this is one of the difficulties of obtaining more detailed secondary evidence about ESP events.

Physiological processes in the subject - and perhaps the experimenter ought to be included too - have been studied for some time in order to find physiological concomitants during ESP processes.

Berger himself was interested in parapsychological research (Roll, 1960) and attempted to develop a theory based on physiological considerations. However Berger apparently did not carry out experiments with simultaneous EEG recordings.

An EEG set-up was donated to the Duke Laboratory about 20 years ago.

Publications have appeared from time to time reporting attempts to find some physiological changes which are connected with parapsychological events (Evans & Osborn, 1952; Wallwork, 1952).

More recently Dean (1961, 1962) has found that certain blood volume changes may be characteristic of successful and unsuccessful

ESP periods, but these changes correspond to levels of success rather than to single hits. Tenny, in an experiment which gave comparatively high ESP scores, also found such overall changes (1962) in GSR and plethysmographic measurements.

Tart's attempt (1962) was perhaps the most sophisticated up to date (see pp. 117-119). Rhine in 1962 in a personal communication with the author, expressed doubt that physiological measurements will ever contribute towards the understanding of parapsychology, but Rhine is sufficiently open minded to support some physiological research in his laboratory.

While the problem of precise physiological measurement to is a point insoluble, there is no a priori reason why physiological concomitants should not be discovered. Since what is recorded in parapsychological experiments, is, in the end, at least partly a physiological response, even a non-physical theory of psi must at some stage allow for the interaction between non-physical and psi processes and the physical processes of the subject and such an interaction could, in principle, lead to special kinds of changes in the physiological processes which can be recorded.

The above suggestion that precision of measurement cannot be improved indefinitely is not only in agreement with micro-physics (Heisenberg, 1955), but can be maintained on the biological level alone. The interference of the measuring

system including the human experimenter, with an organism to be observed, is most obvious when certain measurements terminate the animate existence of that organism.

However there is no reason to assume that the limits of precision in physiological measurements have been reached.

GSR and plethysmograph recordings have led in recent years to encouraging results through the efforts of experimenters like Dean and Tenny.

But the earlier hope that EEG might be an answer to the problem of detecting parapsychological events more directly has not been fulfilled so far. This disappointment in the Berger rhythms was not confined to parapsychologists and was widespread after the early hopes following ^{the} original discoveries. However once the nature of the EEG recordings was better understood, psychologists and physiologists have taken a renewed interest in EEG work and some specific problems have been attacked (Barrett & Herd, 1964).

EEG recordings perhaps more than other physiological measurements, depend a good deal on technique, equipment and on an advanced appreciation of neurological processes which are accompanied by rhythmic changes in the action potential which can be measured and recorded.

On the basis of a theoretical course in EEG interpretation (at Duke University Hospital) and on the basis of practical

measurements at the parapsychology laboratory, and at the Veterans' Hospital (Durham, N.C.) in which the author participated, some recommendations can be made for parapsychological research.

1. It seems most unlikely that progress will be made by simply recording EEG during parapsychological experiments without a sound appreciation of what is recorded. As will be discussed, the recording of EEG during experiments has its own problems.

2. The best and most advanced equipment is just good enough to hope for some results.

3. Interpretation of measurements is difficult and requires careful training.

4. An analyser may help in the interpretation and is a highly desirable addition to a set-up but it cannot replace the trained interpreter of EEG records. An analyser, particularly a digital one, may summarise certain trends more reliably than can be expected from human observers, but information details which could be important in parapsychological research, are lost in this process.

5. An expert technician should be available on a full time or part time permanent basis in order to keep the equipment in optimum working order. To call in such assistance only when a

breakdown has occurred is unsatisfactory and will interfere with research.

For a parapsychological institution EEG work seems only feasible if sufficient finance is available to follow the above recommendations. In this situation team work between a parapsychologist and a neurophysiologist with long experience in EEG work seems highly desirable unless a parapsychologist can rely on experience in both fields. It also seems necessary to plan research along the lines of an overall plan, at least to some extent. A substantial success 10 years after such a project has been started must be considered as fortunate indeed. To work on a day to day basis: "Let us see whether we get something and if not we will try something else" is not likely to succeed.

Some initial steps to find techniques to carry out parapsychological tests during EEG recordings will be included here particularly because few publications about such techniques
27
seem to be available.

27. This work was carried out as a joint effort between Dr. Tenny, who was then a member of the Parapsychological Laboratory, Duke University, and the author. As far as this work was carried out at the Veterans' Hospital, Durham, North Carolina, further assistance was provided through a technician at the Hospital who prepared the subjects (placing of electrodes etc.) for the recordings.

Since speaking or other muscle activities (e.g. pressing a button, but this was not actually tried out) will interfere with the EEG recording, the following technique was tried out and showed signs of practicability. During an EEG recording the subject, who was familiar with ESP cards, was notified by one of the two experimenters (E1) that the other experimenter (E2) either looked at a card or drew it from a pack without looking at it. According to instructions given prior to the EEG recording, the subject was then to 'make up his mind' which call he wanted to make. To ensure a short spontaneous 'response' to the target it may be desirable to restrict the decision time through a mild buzz to a period of say, one second, but this was not attempted during the actual tests. After the decision period E1, who did not handle the target cards, announced slowly the five ESP symbols with a one second pause between each. The technician marked the EEG record when the drawing of the target card was announced and also at the onset of the subject's decision period and at each target symbol when it was spoken by the experimenter. After another period of 5 seconds the subject was asked to name the card which he had selected and this was also marked on the recording.

It was possible to detect the subject's calls from the changes in the EEG recordings without ambiguity. However it is not certain whether all artifacts due to unintentional muscle

movements were eliminated. Nevertheless it appeared that once the procedure developed into a familiar routine, at least gross artifacts in the recording were avoided.

This method appeared to work more satisfactorily than Wallwork's technique (1952) in which the subject called out the symbols. However artifacts due to voice communication may have decreased if a routine had been established.

At this stage no attempts were made to search for physiological differences between hits and misses or to find special hits which might be more particularly due to ESP.

To avoid the necessity of a two experimenter setting and also in order to standardize the situation for each trial it would be desirable to use one and the same tape recording of the necessary announcements, i.e. "select target card - star - circle", etc. Whether the symbols can be presented in the same order throughout depends on whether this would induce the subject to form patterns in the calls (see also pp. 102-107).

In the Parapsychology Laboratory, Duke University, an ESP experiment with GSR and Plethysmograph measurements was conducted in 1962 with Saleh (a member of the laboratory) and the author as joint experimenters.

Saleh was primarily interested in the preferences subjects have in various DT methods and in corresponding ESP results. (Appendix 11, pp. 381-388). The experiment has not been published

but the results were presented at a meeting of the staff of the parapsychology laboratory after this author had left the laboratory and Saleh's report suggested that the ESP hypothesis can be accepted.

However the overall ESP score was not significant but the difference between the scores obtained under the two DT conditions was sufficiently high to reject the null hypothesis.

The present writer was more interested in the physiological measurements and whether it would be possible to find any physiological changes which correspond to ESP hits. For this purpose it was of course important to obtain results which show evidence of ESP. This was the case, but average scores were lower than in Tenny's results.²⁸

However no significant relationship between hits and physiological changes could be discovered. It may be of interest in spite of this, to describe briefly the technique which seemed to have worked satisfactorily and which was similar to that used by Tenny.

Because of the DT methods there was no time sequence of targets. The whole target sequence was ready prior to the experiment.

28. No precise figures can be included here since Tenny's results are not published, as far as is known. (Except for a short report 1962, p. 272) The above remarks are based on personal communications.

The experiment was conducted in two adjoining rooms (R1 and R2). In R2 the subject and one experimenter (E1 - Saleh) were situated. In R2 the second experimenter (E2 - the present writer) operated the recorder. In particular if the subject's GSR changed to an extent that the pen reached the limit of the recording range, E2 balanced the subject's resistance at the recorder. (The recorder had no automatic adjustment). Such changes were marked on the recording paper by E2. During the experiment E1 used a press button switch to mark the simultaneous recording of GSR and blood volume whenever the subjects made a call. E1 also recorded (in writing) the subject's calls on a standard ESP scoring sheet.

The target cards had been prepared by E2 but were unknown to E1. At the completion of the test session E2 marked the target sequence of cards just used on a standard scoring sheet. At this stage E2 did not know the subject's call sequence. The two adjoining rooms were not directly connected by a door and the sound transmission between the two rooms was so weak that E2 could not understand anything the subjects said unless there was deliberate shouting. Hence it is unlikely that E2 included unconsciously any preferential errors when the target sequence was recorded. E2 also numbered the markings from 1 to 25 on the physiological recordings and handed these to E1 who attempted to predict on the basis of changes, at what trials a

hit should have occurred. After these predictions had been completed the target sequence was compared with the call sequence and hits were marked.

In the total experiment 10 test sessions, i.e. 40 ESP runs, were carried out with physiological measurements.

E1 predicted the hits on the basis of the plethysmographic and on the basis of the GSR changes. To see whether these predictions were more or less often correct than could be expected by chance, the total number of correct predictions and incorrect predictions was compared with the total number of correct calls and incorrect calls and the difference was tested by chi square (Appendix 11, pp. 385-386).

The difference was not significant and in no way suggestive in the case of the predictions based on the plethysmograph recording as well as in the case of the predictions based on the GSR recording.

Since the overall ESP score did not reach a significant level the number of correct predictions per ESP run was tabulated for those runs only in which six or more hits had been scored. It was found in a similar analysis as outlined above, that the predictions based on the physiological recordings gave no indication of a significant or suggestive relationship between the predicted hits and the actual hits. Seventeen runs were used in this analysis (Appendix 11 387-388). However,

even if it could be assumed that these differences mean something, it cannot be immediately concluded that physiological changes which varied in relationship to ESP, had been recorded. This would only be possible if a high amount of agreement between say two or three independent judges can be found who make predictions on the basis of the unmarked (as to previous predictions by other judges) recordings of physiological changes.

It seemed to this experimenter that El's choice of certain changes in the physiological records was comparatively arbitrary. El could not formulate a rule that covered all his own selections and while it cannot be said with certainty that no rules for El's selections can be found, this experiment seems to indicate that a much more sophisticated approach (from the physiological point of view) is needed before one can hope to get assistance in parapsychological work through physiological recordings.

These remarks are perhaps too critical because El did not have the sort of ESP result that would seem most suitable for a comparison with physiological changes, and perhaps El's predictions should be considered critically only if a fairly large number of runs with a fairly high number of average hits can be analysed. In particular it seems desirable to go into a more detailed analysis of single subjects which seems difficult in this case, because on the basis of their total runs, the expected frequencies were too small for a normal analysis with chi square.

A GESP test with favourite music targets.

Short reports have appeared previously of two music tests carried out by R. Schulman and R. W. George and which were reported in the J. Parapsychol., 1939, and in the Parapsychol. Bull., 1943, respectively. Although it is impossible to evaluate these early experiments adequately from the short reports which were not presented by the experimenters themselves, it is perhaps surprising that no further investigations into the use of sounds have been made.

The present study however, is not undertaken to compare auditory with visual stimuli, but to explore whether psychological factors which appear to be favourable in music tests will, in fact, aid parapsychological processes. It is difficult to define or even describe the psychological variables in music appreciation adequately, but terms such as "intense, very personal and sometimes highly emotional" would probably find agreement with most who are interested in music at all. It is likely that this sort of experience may be suitable for GESP tests.

In a wider context the exploration into music tests was carried out in the hope of introducing a relatively strong variable (which may be called "the subject's personal involvement in the target variable") which may be comparatively stable over some time and more dominating in its influence on GESP than the sensitive subject-experimenter relationship variable. In other words it seems desirable to see whether music tests may be more

easily repeatable even if the subject-experimenter relationship is less carefully handled. A small improvement in this direction might induce sceptical but interested scientists to try to carry out parapsychological research.

It was also of some interest in connection with other parapsychological research to see whether the provision of special individual targets (favorite pieces of music) would lead to a successful demonstration of GESP.

Procedure.

After nine exploratory runs of 25 trials each under controls which were probably sufficient to rule out sensory cues and after some intermediate investigations into suitable improvements, the following procedure was adopted.

Among volunteer students for parapsychological tests and among Duke psychology students who participated in various tests for which they received credit points (parapsychological tests were acceptable to some psychology instructors), those students were selected who professed to have a strong interest in music and to whom music had a strong personal meaning.

They were asked to select their five favorite pieces of music or five pieces of music which had the strongest personal importance to them. No attempt was made to distinguish between most favorable and most meaningful or important. If subjects had difficulties in selecting five out of a larger number of favorite pieces, they were asked to select those pieces which

they considered to be most contrasting or different.

Subjects were also asked to check whether they could distinguish easily among their selections. Subjects then indicated the most important or representative one minute interval in each of their five selections.

These intervals were played and transferred to magnetic tape. Most subjects furnished their own records.

The five pieces of music were separated on the tape by about 10 seconds of counting from 1 to 10 and by a short spoken title preceding each piece of music.

Depending on previous participation, knowledge, etc. the subjects were introduced to parapsychological research and it was indicated that this music test appeared to have a very good chance of producing high GESP results.

Prior to the first test session the five tape recorded pieces of music were audibly played to each subject. In later sessions the subjects could hear their selections again prior to the experiment if they wanted to do this.

During the test the subject was situated in a room (203C) at the parapsychology laboratory at Duke University and was provided with a standard ESP scoring sheet, a pencil, a Morse key, a wrist-watch (which had a large second hand), and with a sheet of paper on which the five short titles of the music selected were printed in a circular arrangement (Appendix 12, p. 389) to avoid any preferences with respect to the position of any of the five stimuli.

The experimenter was situated in the library of the parapsychology laboratory about three rooms away from the one which the subject occupied. The door of the subject's room was closed. The shortest walking route between this room and the library was about 20 yards. The door of the subject's room was not visible from the library. If the subject pressed the Morse key the experimenter could hear a buzzer in the library. The experimenter could not signal to the subject.

The experimenter could see anybody entering the library and although he had no absolute certainty that the subject remained in his room, an interruption in the subject's buzzing would soon have been noticeable. While it is extremely unlikely that any subject left his room during the test, it is even more unlikely that any subject could have obtained any information about the target without being seen by the experimenter.

During the test the experimenter listened to the pieces of music through dictaphone earphones which were put into the external speaker outlet of the tape recorder. The recorder used was so constructed that the external speaker connection (in this case the earphones) disconnected the internal speaker.

Under these circumstances it was found that none of three persons who checked this could hear anything when they were not more than two feet away from the experimenter and the recorder, and when the music was played at maximum volume.

Prior to each test session comprising usually one run of 25 trials, the experimenter selected the target order from a random number table (Rand Corporation, 1955) by a standard procedure with a dice and coin box (Sanders, 1962, pp.27-28). The relevant numbers determining the starting point in the number table were marked on the experimenter's ESP scoring sheet but not on the subject's sheet.

The subject was instructed to start the first trial by pressing the key twice and it was explained that this double buzz was the experimenter's signal to play and listen to the first one minute music target. During 30 to 60 seconds the subject should attempt to 'feel' what piece of music was playing but if necessary the subject should guess the target if he remained uncertain. The subject could choose to end the playing of the target music after 30 seconds if further time was 'felt' to be of no help.

The subject was instructed to signal the end of the first piece of target music by a single buzz which should be given 30 to 60 seconds after the initial double buzz, i.e. each target was played between 30 and 60 seconds according to the subject's signalling.

It was explained to the subject that it took some time to get the second piece of music ready and the subject was instructed

to wait for 15 seconds after the single buzz before signalling the next double buzz and thereby starting the next trial. The subject was also instructed to write during these 15 seconds, the short name of the 'called' music on the ESP record sheet.

When the instructions were clearly understood, the subject was asked to wait for half a minute before signalling for the first trial, to allow the experimenter to go into the library and to get the first target ready.

The experimenter used the digit of the tape recorder to find the first target (according to the selected random number sequence) and waited for the start signal. The selected music was played and the experimenter listened to it neither with indifference nor with particular excitement as far as this can be judged introspectively. At the beginning of the trial the experimenter wrote down the short name of the music target on his ESP scoring sheet and the corresponding random number was marked off.

As soon as the end of the first trial was signalled the experimenter turned the tape back or forth to find the beginning of the next music target according to the random number table. With the aid of the counter on the tape recorder and markings on the tape spool, this could be achieved within less than 15 seconds.

At the end of each session consisting of 25 trials, the

experimenter walked with his scoring sheet into the room which the subject occupied, and transferred the calls from the subject's scoring sheet to his list of targets. This transfer and the count of hits were carried out in the presence of the subject who checked this operation. There was at no time any ambiguity in this procedure.

The only change in procedure that was introduced for some subjects was to increase the interval between trials from 15 seconds to 20 seconds. This was done whenever in a previous test, a subject scored low on actual hits but high on backward displacement of the order of 1. No changes of any kind were introduced during test sessions.

Prior to the commencement of the formal test (after the pre-test), the number of test runs (consisting of 25 trials each) was fixed as 20. The GESP hypothesis was to be tested on the basis of direct hits only. However, although a positive deviation was expected, the one tail test was not specified.

Other features of interest in the results were to be investigated and analysed if possible but this was not to be taken into account when the GESP hypothesis was tested. These decisions were made on the basis of the results of the pretest.

Results.

In the pre-test the results of 10 runs were to be analysed. However due to technical failure of the buzzer on two occasions only 8 runs were completed and two test sessions consisting of

17 and 11 trials each. Adding these two sessions to one ^{run} +3 trials the pre-test included 9 runs and 3 trials.

In 9 runs and 3 trials $45\frac{3}{5}$ hits would be expected by chance. In the pre-test 61 hits were obtained, and the following values were calculated (Rhine & Pratt, 1957; Owen, 1962, see also Appendix 12, p.390).

TABLE 12

Summary of GESP music results of the pre-test consisting of

9 runs and 3 trials

Deviation	= +13.4
Standard deviation	$2\sqrt{\frac{228}{25}} = 6.05$
Critical ratio	$= \frac{13.4}{6.05} = 2.54$
Probability	= 0.0111
Average score	= 6.69
Total score	= 61
Number of runs	= $9^3/25$

Subjects' direct hits per test session

Subject

A	6 per 17 trials
B	7 per run
B	8 per run
B	4 per 11 trials
C	6 per run
C	5 per run
C	6 per run
D	9 per run
E	6 per run
E	4 per run

The main test consisted of 20 runs, in which 123 hits were obtained and the following values were calculated (Rhine & Pratt, 1937; Owen, 1962).

TABLE 13

Summary of GESP results of the main test consisting of 20 runs

Deviation	= + 23
Standard deviation = $2\sqrt{20}$	= 8.95
Critical ratio = $\frac{23}{8.95}$	= 2.57
Probability	= 0.0102
Average score	= 6.15
Total score	= 123
Number of runs	= 20

TABLE 14

Summary of results of subjects who participated in the main test

Direct hits per run (figures in brackets indicate pre-test
results for complete runs only)

Subject

E	(6) (4) 4, 6, 6,
D	(9) 3, 10, 7,
B	(7) (8) 7, 2, 5, 7, 8,
F	9, 2, 10, 10,
G	5
H	5, 5,
I	5, 7,

To see whether there is any trend in the rate of scoring during the sequence of trials in each test session, the total number of hits per trial 1 to 5, 6 to 10, etc. (for the main test only) were counted (Appendix 12, p. 390) and represented by a graph (Fig. 21). This was also carried out for a sub-group of high and low scoring subjects (Figs. 22, 23). The division into these groups was made on an arbitrary basis. The results of a single subject (F) who contributed most to the positive deviation are shown in a similar graph (Fig. 24).

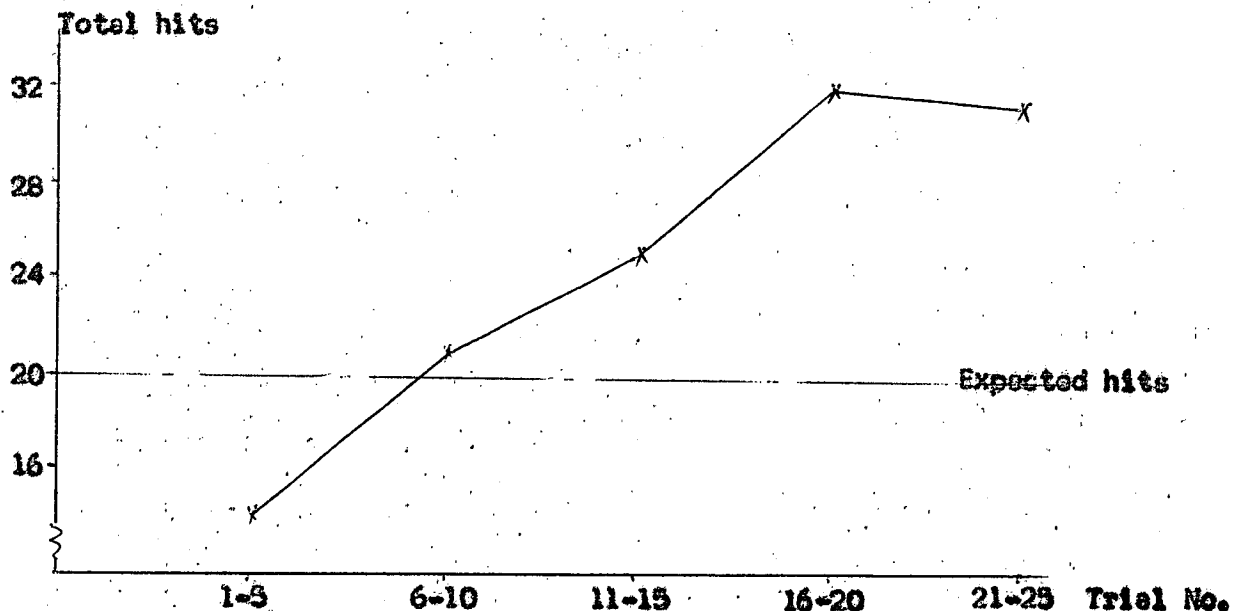


FIG. 21 Total number of hits per interval of five trials in complete main test of 20 GESP music runs.

Seven subjects, total deviation = +23.

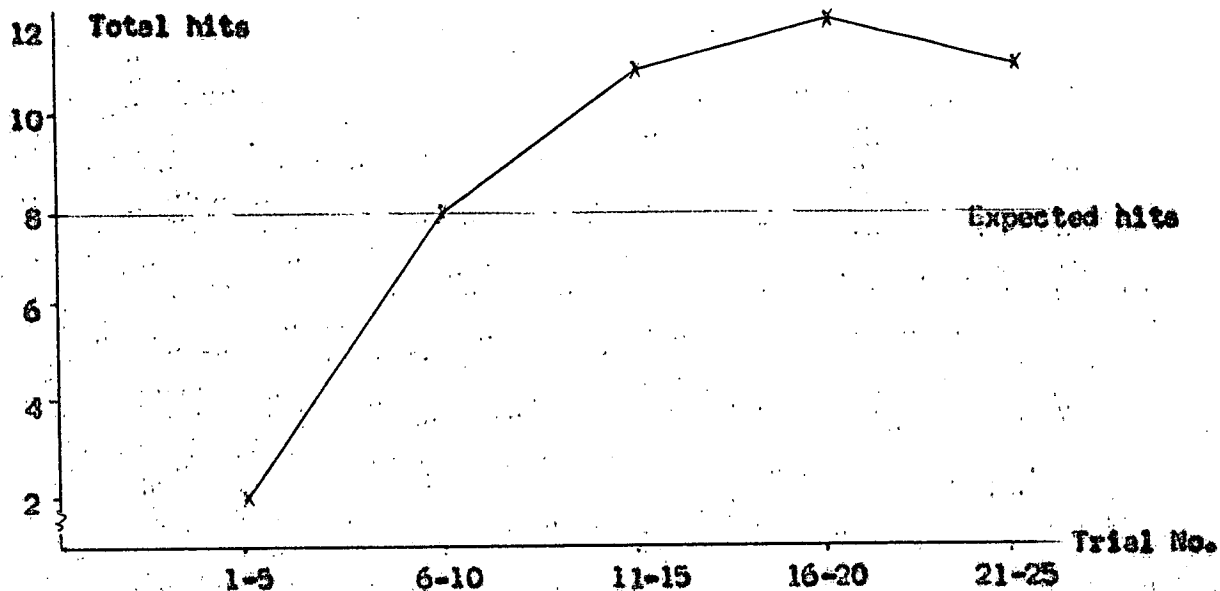


Fig. 22 Total number of hits per interval of five trials in 'low score' part of main test consisting of 8 GESF music runs.
Four subjects, total deviation = +3.

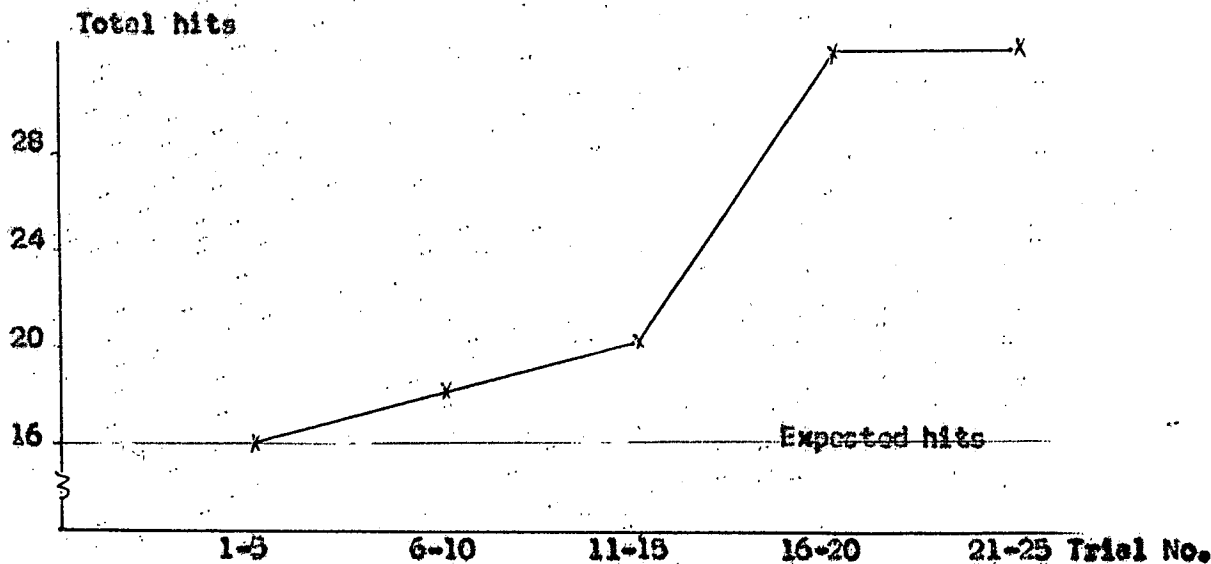


FIG. 23 Total number of hits per interval of five trials in 'high scorer' part of main test consisting of 12 GESF music runs.
Three subjects, total deviation = +20.

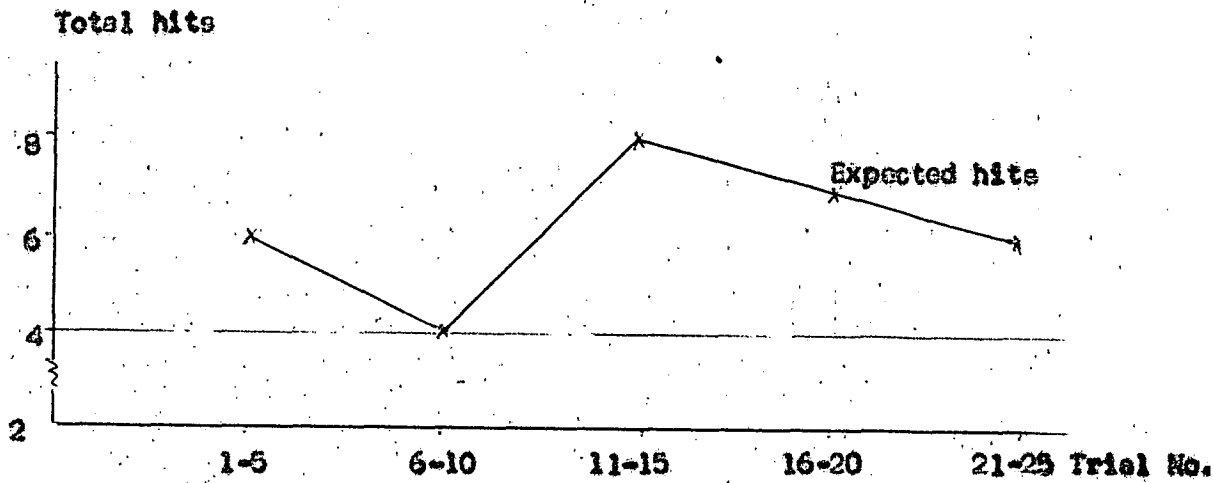


FIG. 24 Total number of hits per interval of five trials for 'highest' scorer* with four GESF music runs in the main test.

One subject (F) total deviation = +11.

Discussion.

The probability that the deviation of +23 occurred by chance is slightly higher than 0.01 which is usually accepted as sufficient to reject the null hypothesis in parapsychology.

It seems justified to reject other counterhypotheses on the basis of the design of the experiment. There is no feasible explanation which could show how some or all subjects obtained information by sensory means.

On the basis of the pre-test the probability of accumulating type 1 errors is at least reduced if not completely eliminated. There are mainly three possibilities left to interpret the deviation.

1. The deviation occurred by chance.
2. The deviation is due to GESP
3. The results are based on fraud.

On a purely subjective basis the experimenter and writer is entitled to reject the third possibility. Although no attempts were made to perform the experiment as a demonstration experiment for hostile sceptics, it can at least be claimed that unintentional errors which would favour the GESP hypothesis are most unlikely to have occurred, particularly since the actual number of trials is fairly small.

The details of the procedure were discussed with Pratt who was then the Deputy Director of the Parapsychology Laboratory at Duke and who was in agreement with the design.

Subjective evidence which would tend to support the GESP hypothesis is that:

1. The experimenter expected a positive deviation after the pre-test results.
2. The pre-test results have a reasonable chance to be also due to GESP.

The main counterhypothesis against GESP in the pre-test could be based on the fact that the experimenter played the tape audibly in the library and that the subject obtained sensory information. It is difficult to rule out this possibility with absolute certainty but on subjective estimations it is unlikely that this happened.

It seems then that one of the aims of the experiment to obtain GESP results has been achieved. If this statement is accepted then the question arises how far the additional aims of the experiment have been reached.

From the individual results of the subjects it is clear that there is no sign of obviously stable positive deviations. The total result rests heavily on subject (F) who had a positive deviation of +11 while the other deviations ranged from 0 to +5.

Perhaps the instability of scores is more noticeable than real if there is any meaning in the high displacement (-1) which occurred for subject (F). When she scored (on direct hits) 2 she had 8 (-1) displacements. There was no general evidence of displacements of this type as the total score reached 104 only when approximately 100 such displacements are expected.

Of course only a larger number of test runs can throw some light on these speculations but at present it can hardly be claimed that the music test gave results which make it a more desirable test than card tests. Considering the results again it is on the other hand, not impossible that the music test has indeed some of the motivational advantages discussed earlier.

The test however also has a number of disadvantages which should be mentioned. It is very slow and as a rule a subject cannot be expected to participate in more than one or two runs during a morning or afternoon. Adding the necessary time for preparation of the individual targets many ESP card runs could be carried out instead of one music run.

There is also no certainty that music plays any direct role in the test at all. The experimenter writes down the target and within approximately 75 seconds the subject writes down the cells.

The possibility that the experimenter could act as an agent is particularly unfortunate since this may be expected to have a bearing on the subject experimenter relationship variable, which is supposed to remain comparatively stable.

To work without an agent is rather difficult because it would be more awkward to describe the target to the subject "as being the music in the earphones which hang on a chair in room X" but it would also require the preparation of complete random sequences of music pieces without the knowledge of the experimenter

and if possible, without anybody else's knowledge. Presumably a juke box might be changed to select pieces in a random sequence but the possibilities with a tape recorder are rather limited.

Within the small number of subject and test runs there is no obvious evidence that the results depended on the experimenter's conscious preferences of music targets although some subjects had selections which were close to the extremes of the experimenter's preference scale.

It is impossible to judge whether the changes of time intervals between trials improved the number of direct hits. However there is evidence that this has happened in previous experiments (Seal & Bateman, 1954).

The graph (Fig. 21) showing the rate of scoring hits in relationship to the sequential position of the scores in one run is not of the more typical U-shape discussed by Pratt (1948). Indeed the scoring rate of hits is quite low in the first five trials per run (Figs. 21, 22) and reaches a maximum in the trials Nos. 16-20.

The subject F with the highest deviation (Fig. 24) differs somewhat from the group results but is - as might be expected on purely logical grounds - more in agreement with the results of the total group than with the U-shaped graph referred to above.

While the speculative nature of this discussion must be emphasised, particularly since the total number of runs is small,

the graphs suggest that GESP processes might have been tapped in a promising sort of way.

The close agreement of Figs. 21 and 22 to psychological learning curves is particularly interesting but it might also suggest to the sceptic that the subjects managed to obtain information by sensory means.

As indicated in the procedure this possibility can be rejected with some confidence. It cannot be categorically stated that it would have been impossible for a subject to leave his room and approach the library between signalling. But one approach to the library was blocked during most of the session through a secretary (unless she is accused of collaboration with the subject) and the second approach was clearly visible to the experimenter over some considerable distance (20' approximately). In either case the subject would have been noticed long before he could have obtained any information from the experimenter's record sheet. Since the name of the target was written down by the experimenter immediately after the subject had signalled the beginning of a trial, the subject could not have signalled and then have run near the library to observe from some distance the movement of a pencil which may convey the word even if the word itself is too far away to be recognised.

Since the confidence level at which the null hypothesis may be rejected is comparatively low (1 per cent approximately) it seems

to this writer more reasonable for the sceptic to assume a No. 1 error instead of counter hypotheses in order to account for the results.

A comparison of the music test with ESP card test seems desirable. But there are some difficulties. An experimenter interested to develop a music test is likely to be biased, and there is some evidence that the experimenter's attitude may influence the card tests (Nash, 1960).

There is also a possibility that a subject's 'correct orientation' towards the music test may be upset through card tests. In particular the wide difference in the speed of the two tests may have a disturbing influence.

For these reasons no card tests were carried out simultaneously with the music test. However when it became evident at the completion of the music test that one subject (F) had comparatively high scores, the experimenter conducted four runs (same number as in the music test) of ESP card tests with her without obtaining a significant or even suggestive deviation.

This does not suggest that this subject can only score successfully in music tests and not in card tests but it does suggest that the subject is not obviously gifted to score a high number of hits under different conditions.

Further experimentation with larger number of runs are desirable and probably justified on the basis of the present study.

It would of course be desirable to find out what role (if any) the music as an auditory stimulus and the experimenter as an agent, play but this is not easy to sort out experimentally. It might be of more immediate interest to see whether different experimenters can obtain similar results with the same group of subjects, i.e. whether the differences between test sessions (same subjects, same experimenter) are of the same order of magnitude as the differences between test sessions (same subjects) when they are carried out by different experimenters.

Summary of experimental reports and future work.

The experiments described here were carried out to see whether certain test conditions and in particular disguised tests, could be developed to demonstrate parapsychological processes.

It seems to this author that this aim has been achieved and particularly the PK results with the automatic apparatus can be regarded as fairly well established.

On rational grounds it can be argued that the test conditions were such, that a repetition under a different experimenter should lead to the successful demonstration of parapsychological events. However this attempt to provide conditions suitable for repetition is not the first one and on the basis of the results presented here it cannot be claimed that a repetition is more easily possible than, say, under normal ESP conditions involving the DT method.

If the rational arguments which led to the assumption of easier repeatability are sound, then the experimental efforts described here can be seen as a first necessary step. Future experiments will show whether it was a step in the right direction or not but it is hoped that the first step was carried out with sufficient care to be able to judge the success or failure of future work.

Ideally such future work should be carried out by different experimenters at different research institutes. However it seems an unrealistic assumption to expect that any suggestions made in

this paper will be followed up by many parapsychologists. Indeed if one or two take these results into consideration when planning their own work the present writer would regard this with some satisfaction.

Perhaps this is too pessimistic a view but it is largely based on the lack of senior students trained in parapsychology. Such students might be prepared to follow up work which has been started elsewhere but the research worker who is starting work in parapsychology and who is already established in another field, usually has more definite plans of his own.

The only way to ensure that the second step of this plan is ever carried out is to continue with one's own work and a few words about such possibilities may be justified.

The PK apparatus is back at the psychology department of the University of Tasmania. For some of the detailed descriptions presented, the apparatus was reconstructed structurally but no attempts have been made so far to put it back into working order.

There is little doubt that this can be done but it may be desirable to carry out certain changes suggested elsewhere which might reduce the necessary maintenance. It should be possible to house the machine in a special test room in the basement of the psychology department in the new building of the Tasmanian University and thus certain control difficulties experienced previously should not arise in future work.

Since it is unlikely that a team approach will develop in

this small department within the near future, some steps should be taken to satisfy the sceptical critics that the experimenter did not interfere with the apparatus or with the recording during a test session and it might be desirable to contact, for instance, G. R. Price, Hansel and Girden to obtain their suggestions.

Obviously it is also of interest to increase the number of variables which are measured in future tests. Personality variables should be considered and reinforcement stimuli perhaps ought to be varied on the basis of a personality assessment. Attitudes towards the experimental situation require special attention. Unfortunately such an expansion is limited by other commitments and will be at best a slow process.

By courtesy of the physics department, University of Tasmania, the essential parts of a microbalance which has been used for PK work previously (see p. 242) were made available to the psychology department in 1964. This apparatus provides the opportunity to influence a physical system by PK where the upper limits of the energy requirements for such an influence are known to be extremely small. However some delicate attachments would have to be constructed to ensure that measurements are not due to ESP.

It seems then that unusual opportunities exist in Tasmania and it is hoped that not all will be wasted.

References

- Abrams, S. I. Paracausal systems and synchronicity. J. Parapsychol., 1958, 22, 299. (brief note only)
- Asher, R. Why are medical journals so dull? Brit. Med. J., 1958, 2, 502.
- Barratt, P. E. H. & Herd, Jean M. Subliminal conditioning of the alpha rhythm. Aust. J. Psychol., 1964, 16, 9-19.
- Battig, W. F. Parsimony in psychology. Psychol. Rep., 1962, 11, 555-572.
- Bender, H. Editorial. Z. f. Parapsychol. u. Grenzgeb. d. Psychol., 1960, 3, 140-148.
- Bindrim, E. A new displacement effect in ESP. J. Parapsychol., 1947, 11, 208-221.
- Bolles, R. C. The difference between statistical hypotheses and scientific hypotheses. Psychol. Rep., 1962, 11, 639-645.

29. Some references included here, e.g. Tart (1962), refer to very brief publications which do not include sufficient detail to enable the reader to form an adequate picture of the research at hand. Such publications have been included (marked 'brief note only') when no other relevant references were known to this author. It is hoped that by pointing to some publication no matter how brief, a first step is taken to secure further information if the reader should find this desirable.

A few references to brief notes in the Parapsychol. Bull. appear in the text but are not listed here if the authors concerned could not be identified. In these cases the necessary details are provided in the text.

- Boring, E. G. The psychology of controversy. Psychol. Rev., 1929, 36, 97-121.
- Boring, E. G. The validation of scientific belief. Proc. Amer. Phil. Soc., 1932, 96, 535-539.
- Boring, E. G. The present status in parapsychology. In Gudas, F. (Ed.) Extrasensory perception. New York : Scribner, 1961.
- Born, M. A statistical interpretation of quantum mechanics. Science, 1955, 122, 675-679.
- Bridgman, P. W. Probability, logic and ESP. Science, 1956, 123, 13-17.
- Bridgman, P. W. Determinism in modern science. In Hook, S. (Ed.) Determinism and freedom. New York University Press, 1958.
- Bridgman, P. W. The way things are. Cambridge, Mass: Harvard University Press, 1959.
- Broad, C. D. Personal identity and survival. London : Society for Psychical Research, 1958.
- Broad, C. D. Discussion. In Scriven, M., Broad, C. D. Pratt, J. G., Burt, C. Physicality and psi : A symposium and forum discussion. J. Parapsychol., 1961, 24, 13-31.
- Broad, C. D. Lectures on psychical research. London : Routledge & Keegan Paul, 1962.
- Broadbent, D. E. Perception and communication. London : Pergamon Press, 1958.
- Brown, S. G. Statistical significance in psychical research. Nature, 1953, 172, 154-156.

- Burt, C. Discussion. In Scriven, M., Broad, C. D., Pratt, J. G.,
Burt, C. Physicality and psi : A symposium and forum discussion.
J. Parapsychol., 1961, 25, 13-31.
- Buytendijk, F. J. J. Mensch und Tier. Hamburg : Rowohlt, 1958.
- Cadoret, R., & Pratt, J. G. The consistent missing effect in ESP.
J. Parapsychol., 1950, 14, 244-256.
- Carlington, W. W. Experiments on the paranormal cognition of drawings
(part I), Proc. Soc. Psych. Res., 1940, 46, 34-151. (a)
- Carlington, W. W. Experiments on the paranormal cognition of drawings
(part II), Proc. Soc. Psych. Res., 1940, 46, 277-334. (b)
- Carpenter, J. C. Brief on 'whole' and 'part' aspects of target cards.
J. Parapsychol., 1963, 27, 268. (brief note only)
- Clark, R. A. The projective measurement of experimentally induced
levels of sexual motivation. J. exp. Psychol., 1952, 44, 391-399.
- Clark, R. A. The effects of sexual motivation on phantasy.
I. McClelland, D. C. (Ed.). Studies in Motivation. New York.
Appleton-Century-Crofts, 1955.
- Cohen, I. B. Orthodoxy and scientific progress. Proc. Amer. Phil. Soc.,
1952, 96, 505-512.
- Cohen, J. Chance, skill, and luck. Middlesex : Penguin Books, 1960.
- Cenani, J. B. Science and common sense. New Haven : Yale University
Press, 1961.
- Cox, W. E. Precognition : an analysis, II. J. Amer. Soc. Psych. Res.,
1956, 50, 99-109.

- Cox, W. E. The influence of 'applied psi' upon the sex of offspring.
J. Soc. Psych. Res., 1957, 39, 65-77.
- Dale, Laura, A., & Woodruff, J. L. The psychokinetic effect : Further
A.S.P.R. experiments. J. Amer. Soc. Psych. Res., 1947, 41, 65-82.
- Dampier, Whetham, W. C. D. A history of science. Cambridge :
University Press, 1929.
- Dean, D. The Plethysmograph as an indicator of ESP. J. Parapsychol.,
1961, 25, 280. (brief note only)
- Dean, D. Further findings . . . using a digital plethysmograph as
an indicator of ESP . . . J. Parapsychol., 1962, 26, 265.
(brief note only)
- Deguisneq, A. The repetitions of the Anderson-White investigation of
teacher-pupil attitudes and clairvoyance test results. Part I.
High school tests. J. Parapsychol., 1959, 23, 196-207.
- Dodds, E. R. Telepathy and clairvoyance in classical antiquity.
J. Parapsychol., 1946, 10, 290-309.
- Dudley, H. C. New Principles in quantum mechanics. New York :
Exposition Press, 1939.
- Eccles, J. The neurophysiological basis of mind. Oxford : Clarendon
Press, 1953.
- English, H. B. & English, Ava C. A comprehensive dictionary of
psychological and psychoanalytical terms. New York : Longmans, Green,
1958.

- Eriksen, C. W. Subception : Fact or artifact? Psychol. Rev., 1956, 63, 74-80. (a)
- Eriksen, C. W. An experimental analysis of subception. Amer. J. Psychol., 1956, 69, 623-634. (b)
- Evans, C. G., & Osborn, E. An experiment in the electroencephalography of medium mistic trance. J. Soc. Psych. Res., 1952, 36, 538-596.
- Eysenck, H. J. Sense and nonsense in psychology. Middlesex : Penguin Books, 1957.
- Fisher, R. A., & Yates, F. Statistical tables (3rd ed.) London : Oliver and Boyd, 1948.
- Fisk, G. W. & West, D. J. ESP tests with erotic symbols. J. Soc. Psych. Res., 1955, 38, 1-7.
- Forwald, H. An approach to instrumental investigation of psychokinesis. J. Parapsychol., 1954, 18, 219-233. (a)
- Forwald, H. PK displacement and air currents. J. Parapsychol., 1954, 18, 41-42. (b)
- Forwald, H. An experimental study suggesting a relationship between psychokinesis and nuclear conditions of matter. J. Parapsychol., 1959, 23, 97-125.
- Garen, D. G. The paradox of pleasure and relativity. New York : Philosophical Library, 1963.
- Gardner, M. Fads and fallacies in the name of science. New York : Dover, 1957.
- George, R. W. In Parapsychol. Bull., 1948, No. 11, 2-3. (brief note only)

- Girden, E. A review of psychokinesis (PK). Psychol. Bull., 1962, 59, 353-388. (a)
- Girden, E. A postscript to "A review of psychokinesis (PK)". Psychol. Bull., 1962, 59, 529-531. (b)
- Goldsmith, Margaret. Franz Anton Mesmer. The history of an idea. London : Barker, 1934.
- Guilford, J. P. Psychometric methods. (2nd ed.) New York : McGraw Hill, 1954.
- Guilford, J. P. Fundamental statistics in psychology and education. (3rd ed.) New York : McGraw Hill, 1956.
- Gurney, E., Myers, F. W. H., & Podmore, F. Phantasma of the living. London : Trubner, 1886.
- Hansel, C. E. M. A critical analysis of the Pearce-Pratt experiment. J. Parapsychol., 1961, 25, 87-91. (a)
- Hansel, C. E. M. A critical analysis of the Pratt-Woodruff experiment. J. Parapsychol., 1961, 25, 99-113. (b)
- Heisenberg, W. Das Naturbild der Heutigen Physik. Hamburg : Rowohlt, 1955.
- Hilton, H., Baer, G., & Rhine, J. B. A comparison of three sizes of dice in PK tests. J. Parapsychol., 1943, 7, 172-190.
- Hoyle, F. Man and materialism. London : Allen & Unwin, 1957.
- Hull, C. L. A behaviour system. New Haven : Yale University Press, 1952.
- Humphrey, Betty M. ESP subjects rated by two measures of personality. J. Parapsychol., 1949, 13, 274-291.
- Humphrey, Betty M. ESP score level predicted by a combination of measures of personality. J. Parapsychol., 1950, 14, 193-206.

- Humphrey, Betty M. The relation to some personality ratings to ESP scores. J. Soc. Psych. Res., 1951, 36, 453-466.
- Humphrey, Betty M., & Rhine, J. B. PK tests with six, twelve, and twenty-four dice per throw. J. Parapsychol., 1944, 8, 139-157.
- Humphrey, Betty M., & Rhine, J. B. PK tests with two sizes of dice mechanically thrown. J. Parapsychol., 1945, 9, 124-132.
- Hyslop, J. H. Enigmas of psychical research. London : Putnam, 1906.
- Jessup, M. K. The case for the UFO. London : Arco, 1935.
- Jordan, P. Atomic physics and parapsychology. Newsletter Parapsychol. Found., 1955, 2, 3-7.
- Jung, C. G. (1936) The archetypes and the collective unconscious. In The Collected works by C. G. Jung. (Vol. 9, part 1) London : Routledge & Keegan Paul, 1959.
- Jung, C. G. Ein Brief zur Frage der Synchronizität. Z. f. Parapsychol. u. Grenzgeb. d. Psychol., 1961, 3, 1-9.
- Kahn, S. D. The enigma of psi : A challenge for scientific method. J. Amer. Soc. Psych. Res., 1962, 56, 114-124.
- Kahn, S. D. & Neisser, U. A mechanical scoring technique for testing GESP. J. Parapsychol., 1949, 13, 177-185.
- Keil, H. H. J. Some adaptations of standard equipment for automatic psychological recordings. Aust. J. Psychol., 1961, 13, 99-105.
- Keil, H. H. J. An apparatus for very short projection times. Aust. J. Psychol., 1962, 14, 39-40.
- Kellog, W. N. Porpoises and sonar. University of Chicago Press, 1961.

Kennedy, J. L. The visual cues from the backs of ESP cards.

J. of Psychol., 1938, 6, 149-153.

Kennedy, J. L. An evaluation of extra-sensory perception. Proc.

Amer. Phil. Soc., 1952, 96, 513-518.

Konecni, E. B. Bioastronautics review - 1963, NASA Headquarters, Washington, D.C., 1963.

Kouwenhoven, W. B. Effects of electricity on the human body.

Electrical Engineering, 1949, 68, 199-203.

Kuder, F. G. Kuder Preference record occupational, F or D research handbook. (2nd ed.) Chicago : Science Research Associates, 1957.

Landé, A. The case for indeterminism. In Hook, S. (Ed.)

Determinism and freedom. New York University Press, 1958.

Landé, A. Quantum mechanics, from duality to unity. American Scientist, 1959, 47, 341-349.

Layton, W. L. (Ed.) The strong vocational interest blank : research and uses. Minneapolis : University of Minnesota Press, 1960.

Leroy, O. La levitation. Paris, Valois, 1928.

McConnell, J. V., Cutler, R. L., & McNeil, E. B. Subliminal stimulation :

An overview. Amer. Psychologist, 1959, 13, 229-242.

McConnell, R. A. Physical or non-physical ? J. Parapsychol., 1947, 11, 111-117.

McConnell, R. A. Why throw dice? J. Parapsychol., 1952, 16, 187-191.

- McConnell, R. A. The nature of the laboratory evidence for extrasensory perception. In Holstenholme, G. E. W., & Miller, Elaine, C. P. (Eds.) Ciba Foundation symposium on extrasensory perception. London : Churchill, 1956.
- McConnell, R. A. Scaled measurements in psi research. J. Parapsychol. 1958, 22, 237-245.
- McConnell, R. A., Snowden, R. J., & Powell, K. F. Wishing with dice. J. exp. Psychol., 1955, 50, 269-275.
- McDougall, W. Editorial introduction. J. Parapsychol., 1937, 1, 1-9.
- McElroy, W. A., & Brown, W. R. K. Electric shocks for errors in ESP card tests. J. Parapsychol., 1950, 14, 257-266.
- McGeoch, J. A. The psychology of human learning. New York : Longmans, Green, 1942.
- McGuigan, F. J. Experimental psychology. Englewood-Cliffs, N.J. : Prentice-Hall, 1960.
- McNemar, Q. Psychological statistics. (2nd ed.) New York : Wiley, 1955.
- McWhirter, K. G. Comments on Mr. Cox's essay. J. Soc. Psych. Res., 1957, 39, 77-78.
- Mangan, G. L. How legitimate are the claims for ESP? Aust. J. Psychol., 1959, 11, 121-125.
- Margenau, H. A principle of resonance. Newsletter Parapsychol Found., 1957, 4, 3-6.
- Meehl, E., & Scriven, M. Compatibility of science and ESP. Science, 1956, 123, 14-15.

- Mitchell, A. J. M., & Fisk, G. W. Clock cards and differential scoring techniques. J. Parapsychol., 1954, 18, 153-164.
- Morgan, C. L. An introduction to comparative psychology. New York : Scribner, 1901.
- Murphy, G. Psychical research and the mind-body relation. J. Amer. Soc. Psych. Res., 1946, 4C, 189-207.
- Murphy, G. Progress in parapsychology. J. Parapsychol., 1958, 22, 229-236.
- Murphy, G. Trends in the study of extrasensory perception. Amer. Psychologist, 1958, 13, 69-76.
- Murphy, G. Experimental objectives of the Anderson-White work with school children. J. Parapsychol., 1959, 23, 133-134.
- Murphy, G. Challenge of psychical research. New York : Harper, 1961.
- Myers, F. W. H. Human personality and its survival of bodily death. (Vol. 2), New York : Longmans, Green, 1903.
- Nash, C. B. The PK mechanism. J. Soc. Psych. Res., 1955, 38, 8-11.
- Nash, C. B. Interpretation of the results of Forwald's experiments in placement PK. J. Parapsychol., 1956, 20, 53-55.
- Nash, C. B. The effect of subject-experimenter attitudes on clairvoyance scores. J. Parapsychol., 1960, 24, 189-198.
- Nash, C. B. A comparison of responses to extrasensory and subliminal targets. J. Parapsychol., 1961, 25, 278-279.
- Nash, C. B., & Durkin, M. G. Terminal salience with multiple digit targets. J. Parapsychol., 1959, 23, 49-53.

- Onetto, B. To the editors. J. Parapsychol., 1962, 26, 57-59.
- Oppenheimer, J. R. Science and the common understanding. London : Oxford University Press, 1954.
- Oram, A. T. An experiment with random numbers. J. Soc. Psych. Res., 1954, 37, 369-377.
- Osis, K. A test of the relationship between ESP and PK. J. Parapsychol., 1953, 17, 298-309.
- Osis, K. The distance problem. J. Parapsychol., 1959, 23, 290.
- Osis, K., & Pienaar, D. C. ESP over a distance of seventy-five hundred miles. J. Parapsychol., 1956, 20, 229-232.
- Owen, D. B. Handbook of statistical tables. London : Addison-Wesley, 1962.
- Picken, D. A. The effects of electricity on human beings. The Institution of Electrical Engineers Abstracts, 1961, No. 3492, 19-20.
- Pratt, J. G. Approaches to the investigation of ESP. J. Amer. Soc. Psych. Res., 1942, 36, 122-139.
- Pratt, J. G. The meaning of performance curves in ESP and PK test data. J. Parapsychol., 1949, 13, 9-22.
- Pratt, J. G. The variance for multiple-calling ESP data. J. Parapsychol., 1954, 18, 37-40.
- Pratt, J. G. Discussion. In Scriven, M., Broad, C.P., Pratt, J. G., Burt, C. Physicality and psi : a symposium and forum discussion. J. Parapsychol., 1961, 25, 13-31.
- Pratt, J. G. A visit to parapsychological research centres in Russia and Czechoslovakia. J. Parapsychol., 1962, 26, 269-270.

Pratt, J. G. The Girden-Murphy papers on PK. J. Parapsychol., 1963, 27, 199-209.

Pratt, J. G., & Forwald, H. Confirmation of the PK placement effect. J. Parapsychol., 1958, 22, 1-19.

Pratt, J. G., & Woodruff, J. L. Size of stimulus symbols in extra sensory perception. J. Parapsychol., 1939, 3, 121-158

Pratt, J. G., & Woodruff, J. L. Refutation of Hansel's allegation concerning the Pratt-Woodruff series. J. Parapsychol., 1961, 25, 114-129

Price, G. R. Science and the supernatural. Science, 1955, 122, 359-367.

Price, G. R. Where is the definite experiment? Science, 1956, 123, 17-18.

Price, H. H. Apparitions : two theories. J. Parapsychol., 1960, 24, 110-128.

Pronko, N. H. A classroom demonstration of "extrasensory perception" Psychol. Record, 1961, 11, 423-425.

Rand Corporation. A million random digits with 1000000 normal deviates. Glencoe, Ill., Free Press, 1955.

Reeves, Margaret, P., & Rhine, J. B. The PK effect : II. A study in declines. J. Parapsychol., 1943, 7, 76-93.

Rhine, J. B. Extra-sensory perception. London : Faber, 1935.

Rhine, J. B. The reach of the mind. Middlesex : Penguin Books, 1954. (a)

Rhine, J. B. New world of the mind. London : Faber, 1954. (b)

Rhine, J. B. The experiment should fit the hypothesis. Science, 1956, 123, 19.

Rhine, J. B. Why national defence overlooks parapsychology.

J. Parapsychol., 1957, 21, 243-258.

- Rhine, J. B. On the nature and consequences of the unconsciousness of psi. J. Parapsychol., 1958, 22, 175-186.
- Rhine, J. B. Some avoidable misconceptions in parapsychology. J. Parapsychol., 1959, 30-43.
- Rhine, J. B. Extra-sensory perception : a review. In Gidas, F. (Ed.), Extra-sensory perception. New York : Scribner, 1961.
- Rhine, J. B. The precognition of computer numbers in a public test. J. Parapsychol., 1962, 26, 244-251. (a)
- Rhine, J. B. The shifting scene in parapsychology. J. Parapsychol., 1962, 26, 293-307. (b)
- Rhine, J. B. Parapsychology at Duke : retrospect and prospect. Parapsychol. Bull., 1963, No. 65, 1-6.
- Rhine, J. B., & Humphrey, Betty M. A transoceanic ESP experiment. J. Parapsychol., 1942, 6, 52-74.
- Rhine, J. B., & Pratt, J. G. A review of the Pearce-Pratt distance series of ESP tests. J. Parapsychol., 1954, 18, 165-177.
- Rhine, J. B., & Pratt, J. G. Parapsychology : Frontier science of the mind. Springfield, Ill. : Charles C. Thomas, 1957.
- Rhine, J. B., & Pratt, J. G. A reply to the Hansel critique of the Pearce-Pratt series. J. Parapsychol., 1961, 25, 92-98.
- Rhine, J. B., Pratt, J. G., Smith, B. M., Stuart, C. F., & Greenwood, J. A. Extrasensory perception after sixty years. New York : Holt, 1940.
- Rokeach, M. The open and closed mind. New York : Basic Books, 1960.
- Roll, W. G. Psyche by Hans Berger. J. Parapsychol., 1960, 24, 142-148.

- Roll, W. G. The problem of precognition. J. Soc. Psych. Res., 1961 41, 115-128.
- Rosenthal, R., & Gaito, J. The interpretation of levels of significance by psychological researchers. J. Psychol., 1963, 55, 33-38.
- Ruppelt, E. J. The report on unidentified flying objects. London : Gollancz, 1956.
- Rýzl, M. A brief report of a journey to the USSR. Parapsychol. Bull., 1961, 59, 1-3. (b)
- Rýzl, M. Research on telepathy in Soviet Russia. J. Parapsychol., 1961, 25, 75-85. (a)
- Rýzl, M. Biologicheskaya Radiosvysz (biological radio) by B. B. Kozhinsky, J. Parapsychol., 1962, 26, 221-226.
- Rýzl, M. Long-distance suggestion by L. L. Vasiliev. Moscow : Gospolitizdat, 1962. J. Parapsychol., 1963, 27, 50-55.
- Rýzl, M., & Pratt, J. G. Confirmation of ESP performance in a hypnotically prepared subject. J. Parapsychol., 1962, 26, 237-243.
- Rýzl, M. & Pratt, J. G. The focusing of ESP upon particular targets. J. Parapsychol., 1963, 27, 227-241. (c)
- Rýzl, M. & Pratt, J. G. A further confirmation of stabilized ESP performance in a selected subject. J. Parapsychol., 1963, 27, 73-83. (a)
- Rýzl, M. & Pratt, J. G. A repeated-calling ESP test with sealed cards. J. Parapsychol., 1963, 27, 161-174. (b)

- Ryzi, M., & Ryzlova, Jirina. A case of high-scoring ESP performance in the hypnotic state. J. Parapsychol., 1962, 26, 153-171.
- Sanders, M. S. Comparison of verbal and written responses in a precognition experiment. J. Parapsychol., 1962, 26, 23-34.
- Savage, L. J. The foundation of statistics. New York : Wiley, 1954.
- Schlesinger, G. The method in the physical sciences. London : Routledge & Keegan Paul, 1963.
- Schmeidler, Gertrude R. Separating the sheep from the goats. J. Amer. Soc. Psych. Res., 1945, 39, 47-50.
- Schmeidler, Gertrude, R. Research projects in parapsychology. J. Parapsychol., 1948, 12, 107-113.
- Schmeidler, Gertrude, R. Personality correlates of ESP as shown by Rorschach studies. J. Parapsychol., 1949, 13, 23-30.
- Schmeidler, Gertrude, R. Exploring the parameters of research variables. J. Parapsychol., 1959, 23, 238-250.
- Schmeidler, Gertrude R., & McConnell, R. A. ESP and personality patterns. New Haven : Yale University Press, 1958.
- Schulman, R. An experiment in ESP with sounds as stimuli. J. Parapsychol. 1938, 2, 322-325.
- Scriven, M. Some theoretical possibilities in psi research. J. Soc. Psych. Res., 1957, 39, 78-83.
- Scriven, M. New frontiers of the brain. J. Parapsychol., 1961 25, 308-318.
- Scriven, M. Scientific method : the foundation of psychology. In Calvin, A. D. (Ed.) Psychology. Boston : Allyn & Bacon, 1961.

Scriven, M. To the editors. J. Parapsychol., 1962, 26, 56-57. (a)

Scriven, M. To the editors. J. Parapsychol., 1962, 26, 131-133. (b)

Sheffield, F., & Kaufman, R. S. Letters to J. B. Rhine. In Editorial.

A PK experiment at Yale starts a controversy. J. Amer. Soc. Psych. Res. 1952, 46, 111-117.

Sherrington, C. Man on his nature. Middlesex : Penguin Books, 1955.

Smith, R. S., Dagle, E. F., Hill, M. D., & Mott-Smith, J. Testing for extrasensory perception with a machine. Research Report, 1963, AFCRL - 63 - 141, Air Force Cambridge Research Laboratories, Hansom Field, Mass.

Smythies, E. A. A case of levitation in Nepal. J. Soc. Psych. Res., 1951, 36, 415-426.

Smythies, J. R. The extension of mind : a new theoretical basis for psi phenomena. J. Soc. Psych. Res., 1951, 36, 477-502.

Soal, S. G., & Bateman, F. Modern experiments in telepathy. New Haven : Yale University Press, 1954.

Stevens, C. C. Parapsychology and modern science. J. Parapsychol., 1950, 14, 127-139.

Stevenson, I. To the editors. J. Parapsychol., 1962, 26, 59-64.

Stuart, C. E., Humphrey, Betty M., Smith, B. M., McMahon, Elizabeth. Personality measurements and ESP tests with cards and drawings. J. Parapsychol., 1947, 11, 118-146.

Sullivan, J. W. N. The limitation of science. New York : Mentor Book, 1949.

Tart, C. T. An experiment comparing the autonomic responses of the percipient ... J. Parapsychol., 1962, 26, 265. (brief note only)

- Tenhaeff, W. H. C. Aussergewöhnliche Heilkräfte. Freiburg ; Walter Verlag, 1957.
- Tenny, K. Physiological responses during an ESP test. J. Parapsychol., 1962, 26, 272. (brief note only)
- Thouless, R. H. Some experiments on PK effects in coin spinning. J. Parapsychol., 1945, 9, 169-175.
- Thouless, R. H. & Weisner, B. P. The psi processes in normal and "paranormal" psychology. J. Parapsychol., 1948, 12, 192-212.
- Tornier, E. Die Arbeitshypothese 'Antizufallswahrscheinlichkeit' - ihr Ursprung und ihre Grenzen. Z. f. Parapsychol. u. Grenzgeb. d. Psychol., 1960, 3, 90-119.
- Tornier, E. In written communications with Bender, H. Institut für Grenzgebiete der Psychologie und Psychohygiene, Freiburg, Germany, and with the present writer, 1963.
- Tyrrell, G. N. M. Perspective in psychical research. J. Amer. Soc. Psych. Res., 1946, 40, 229-240.
- Tyrrell, G. N. M. Science and psychical phenomena & apparitions. New York : University Books, 1961.
- Van Busschbach, J. G. An investigation of ESP between teacher and pupils in American schools. J. Parapsychol., 1956, 20, 71-80.
- Van Busschbach, J. G. An investigation of ESP in first and second grades in American schools. J. Parapsychol., 1961, 25, 161-174.
- Wallwork, S. C. ESP experiments with simultaneous electroencephalographic recordings. J. Soc. Psych. Res., 1952, 36, 697-701.

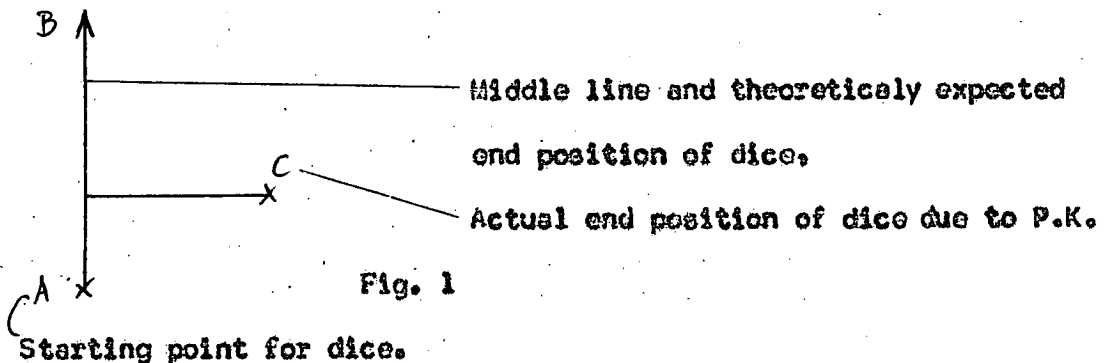
- Walter, W. G. The living brain. London : Duckworth, 1953.
- Walter, W. G. The neurophysiological aspects of hallucinations and illusory experiences. London : Society for Psychical Research, 1960.
- Wassermann, G. D. An outline of a field theory of organismic form and behaviour. In Wolstenholme, G. E. W., & Miller, Elaine, C. P. (Eds.) Ciba Foundation Symposium on extrasensory perception. London : Churchill, 1956.
- Wechsler, D. The measurement and appraisal of adult intelligence. (4th ed.) Baltimore : Williams & Wilkins, 1958.
- Werthmann, H. V. Signifikante Effekte in Zufallsverteilungen als Problem der quantitativen Forschung in der Parapsychologie. Z. f. Parapsychol. u. Grenzgeb. d. Psychol., 1961, 5, 48-69.
- West, D. J. The strength and weakness of the available evidence for extrasensory perception. In Wolstenholme, G. E. W., & Miller, Elaine C. (Eds.) Ciba Foundation Symposium on extrasensory perception. London : Churchill, 1956.
- Whittlesey, J. R. B., Greenwood, J. A. Some comments apropos of pooling. J. Parapsychol., 1959, 23, 135-137.
- Wiener, N. The human use of human beings. London : Eyre & Spottiswoode, 1954.
- Wolstenholme, G. E. W., & Miller, Elaine, C. P. (Eds.) Ciba Foundation symposium on extrasensory perception. London : Churchill, 1956.
- Woodruff, J. L. To the editors. J. Parapsychol., 1961, 25, 266-268.
- Woodruff, J. L. To the editors. J. Parapsychol., 1962, 26, 133-134.
- Zweig, S. Die Heilung durch den Geist. Wien : Reicher, 1936.

Appendix 1

The copy of the letter included here had been incorporated in an experimental report which was submitted for examination purposes previously.

September, 1959.

"Today I would like you or Dr. Pratt or anybody else interested to consider the following argument in connection with Mr. Forwald's results. Mr. Forwald calculated an actual side force. Without knowing his arguments which led to his calculations in detail it appears to me that his fundamental ideas must be as follows:



If we call the perpendicular distance from C to AB ' d ' then I think Mr. Forwald assumes that a force acted on the dice which is roughly equivalent to shifting a mass representing the dice over a distance ' d '. I would disagree with this on the

following grounds. From A to C the dice roll through a number of positions each of which is equal or nearly equal to an unstable equilibrium. In each of these positions only a minute force (of which the magnitude is quite uncertain, as far as I can see), would be necessary to shift the dice from left to right or vice versa.

To consider one of these positions should be sufficient to describe the sequence.

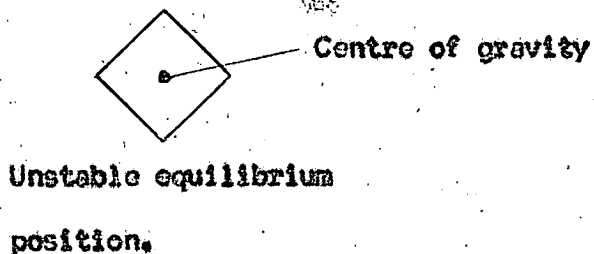


Fig. 2

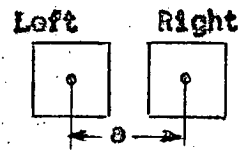
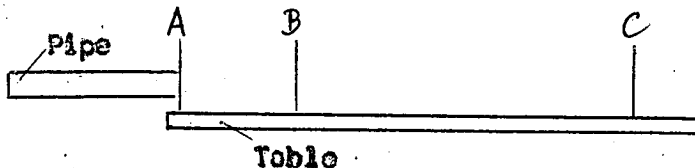


Fig. 3

A small force can shift dice from the unstable equilibrium position to right or left,

i.e. through a small force of uncertain magnitude we obtain a shift of a distance 's'. This force cannot be compared with the one necessary to shift a certain mass (dice) represented by its centre of gravity, the same distance 's'. It seems to me that the total distance 'd', is merely made up from several distances of type 's', where in each case only a small force of uncertain magnitude is acting. Hence the total force producing the deviation from the centre line (distance 'd') is small and of uncertain magnitude.

It may be argued that this (i.e. my own view) is not correct when the dice have developed a certain momentum prior to the section where they are under P.K. influence, i.e. not a standing but a moving dice with a certain directional force is changing its course. It is true, I think, that from a certain speed onwards my first argument does not hold. i.e. if the dice were thrown with a considerable speed in one direction, for instance, mechanically through a pipe which is in a horizontal position on a horizontal table then it seems likely that



there will be a certain distance AB over which unstable equilibrium positions have little influence on the course of the dice. However, as the dice run slower and slower, the unstable equilibrium positions - and possibly P.K. - will determine more and more the final resting points. The material and surface of the dice and table may well have something to do with where 'B' starts.

Apart from this general consideration it seems to me that Mr. Forwald's dice falling through the air had little directional force in a horizontal direction and I believe that the AB section is very small in comparison to the BC section; or in other words, that the change in the course of the dice can be

attributed to P.K. acting on the unstable equilibrium positions, i.e. the actual magnitude (in Dyn) of the side force remains uncertain."

In a reply to this letter Pratt agreed with the arguments set out by this writer and also mentioned that Nash (1956) had reached a similar position.

Appendix 2

Chi square analysis to test the difference between the observed frequencies of end calls (and non-end calls) from one subject in the Smith, et al. (1963) experiment and the expected frequencies.

Symbols called ranged from 1 to 10 and response buttons for the symbols were presented in this order, 1, 2, 3, ..., 10.

10 = \emptyset	fo	fe	total
end calls, i.e. symbols 1, \emptyset	9	20	29
non-end calls, i.e. symbols, 2,3,...,9.	91	80	171
total	100	100	200

$$\text{chi square} = \frac{11^2}{20} + \frac{11^2}{80}$$

$$= 6.05 + 1.51$$

$$= 7.56$$

$$p < 0.01 ; df = 1$$

Appendix 3

Research possibilities in PK

Pages 311 to 330 included here had been incorporated in an unpublished experimental report which was submitted for examination purposes previously.

What I am trying to do now is to consider one branch of psi only: P.K. and to consider the research possibilities and particularly new possibilities in more detail. I am also attempting to put forward some goals and broader classifications which will group some of these possibilities.

These possibilities are based on evidence and speculation. The evidence may be classified as follows:

Primary Evidence ('A' evidence).

- | | | | |
|------------------|---|------------|-------------------------------|
| 1. E.S.P. exists | } | psi exists | based on statistical evidence |
| 2. P.K. exists | | | with high significance |

Satisfactory Secondary Evidence ('B' evidence).

1. Decline of P.K. in a continuous experimental series.
2. Large differences in the significance of P.K. results occur.
3. Subjects are not usually aware of differences in the success or otherwise of P.K. scores.
4. Certain positive findings such as : Subject 'X' had significant results (at 'Y' level of significance) over a distance 'C'.
5. P.K. occurs in situations where the magnitudes of forces that could produce the scored change may be small, perhaps extremely small. (24)

Doubtful Secondary Evidence ('C' evidence) (see also page 12).

- | | | |
|-----------------------------|---|--------------------|
| 1. P.K. space independent.) | } | P.K. non-physical. |
| 2. P.K. time independent.) | | |

3. P.K. depends on subject's motivation.
4. P.K. depends on experimenter-subject relationship.
5. P.K. depends on subject's belief in P.K.
6. P.K. depends on other personality factors, mental or physiological.

Doubtful Secondary Evidence Derived from Single Cases ('D' evidence).

1. Under high tension or motivation the subject increases the P.K. output, but the 'direction' (42) is often wrong.

All evidence is derived from experimental situations which can be described as similar to scientific experimental situations. No evidence is included from experiments with mediums, i.e. from situations which are too unfamiliar to be accepted as evidence when a good deal of fraud was discovered. This does not necessarily mean that some of the evidence could not be very valuable in guiding research. More about this will be mentioned on page 39.

Speculation 1 (based on 'B' 2). ϕ

- ('S' 1) P.K. ability is an ability which exists in some members of the total population only. Sir Charles Sherrington (32) mentioned 'tasters' as an example of a hereditary ability confined to a limited number of human beings.

Speculation 2a and 2b (based on 'B' 2).

P.K. ability either exists in all and has been lost (beyond recapture) by some, or P.K. ability does not exist at birth but some acquire it in their development at a certain stage.

Past this stage P.K. ability cannot be acquired any longer.

'S'1, 'S' 2a and 'S' 2b suggest the same experimental situations: under optimal conditions only some subjects are able to score. 'S'1 stresses hereditary factors while 'S'2a and 'S' 2b stresses the environmental and developmental factors.

Speculation 3 (based on 'B' 2).

('S'3) P.K. ability exists in all and can operate under optimal conditions which have not been found yet.

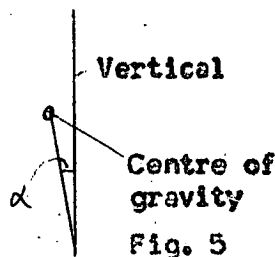
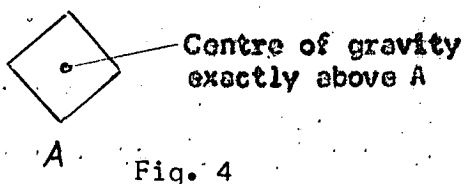
Speculation 4 (based on B1, B2, B4 and B5).

('S'4) Assuming that 'S' 3 is true and that conditions are optimal or that this is not necessary, the large difference in P.K. scores could occur if the 'direction' of P.K. is wrong. By correct direction is meant: P.K. acts correctly in the prescribed way. For instance, if the number six of a die is wished for and P.K. acts towards this goal. But it could be that P.K. does in fact act but towards another goal, say the number one. The evidence (B4 evidence) that this can happen in a systematic way (i.e. for the opposite or for an adjacent number) is strong (43, 44). Therefore it is not unreasonable to assume that P.K. may be active all the time but simply non-directive when no significant scores are obtained. This new term 'non-directive' may sound like an excuse for non-significant results. However, this speculation has certain experimental

possibilities which seem to justify this new term.

Speculation 3 (based on B1 and E6)

('S'3) Even if P.K. acts and the direction is correct it may be that not all experimental runs have conditions where P.K. can influence the outcome or final results. This can be made clearer if we consider a die in a true unstable equilibrium position.



In this position a minute force could push the dice either way. If we have a true unstable equilibrium position, or a position very close to it then it may be assumed that P.K. can produce a scoreable result.

But if we assume that angle α is the maximum angle which P.K. could overcome then in an experimental run in which α is never reached (i.e. the actual angle is larger) then it is possible that we do not score P.K. even if it was acting in the correct direction. α If we represent the die by the centre of gravity 'C' then α is the angle which could be overcome (Fig. 5). For example if the desired direction is towards the right and if 'C' is in a position at an angle smaller or equal to α to the left then P.K. could overcome this angle and shift C to the right.

Fig. 6



movement due to
P.K.

But if during one run, due to the initial position of the die, and its initial momentum, all important (i.e. those that decide the right or wrong direction), unstable equilibrium positions have an angle which is larger than α at the wrong side, then even if P.K. produces a maximum change, i.e. shifts the centre of gravity through the angle α in the correct direction, then this will not be sufficient for a correct score.

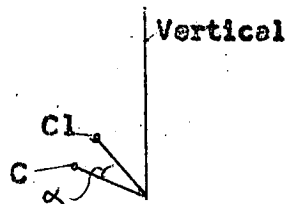


Fig. 7

After C is shifted to Cl the final position will still be to the left, i.e. to the wrong side.

This speculation is a fairly clear one for dice and has therefore some straight-forward possibilities in dice experiments. However remembering that α would have to be extremely small if the force necessary to overcome it shall be anywhere near the Heisenberg uncertainty limit, it seems possible that the same argument may apply to other situations where P.K. can be scored. For instance, a ball which can move to the right or left lane from a single line is, even if accurately manufactured (steel ball for ball bearings) not exactly round, and it may be argued that the ball goes through unstable equilibrium positions too.

* This may be a true unstable equilibrium position as in Fig. 4 or an apparent unstable equilibrium position as in Fig. 5 and Fig. 6.

Speculation 6

('S'6) The process of P.K. is accompanied by physiological concomitants in the subject.

All speculations are developed as extreme cases. They may all be true in a limited way.

It would not be unusual to call these speculations hypotheses and there is some justification for this, since they can be tested. However, at present we can only expect that some of the speculations will lead to 'B'4 evidence, i.e. their falsehood cannot be proved.

Before I come to the possibilities I would like to put forward some goals or one main goal and several sub-goals first.

The main goal is to find a theory that will explain the evidence in a satisfactory manner.

First sub-goal: To provide satisfactory secondary evidence in addition to such evidence already mentioned, i.e. to supply more facts which the theory is supposed to explain, or to find lawful relationships which may be regarded as components of the final theory.

Second sub-goal: To assist in the establishment of satisfactory secondary evidence. The reliability which is low (due to 'B'1, 'B'2, 'B'3, 'B'5, and 'B'6) should be increased. This seems to be the key to advancement towards the main goal.

The six speculations and the evidence suggests the following possibilities to advance P.K. research. Some of the possibilities are obvious and others have been stated elsewhere, however it seems desirable to list them all.

As indicated earlier in this paper I do not suggest in any way that this list is likely to be complete or even nearly so.

Possibilities.

P.1 Improvement of the experimental design by allowing for a better statistical analysis. ('A' evidence).^{*}

P.2 To investigate hereditary factors in connection with successful P.K. scores. ('S'1).

- (a) Intelligence tests based on Piaget rather than Wechsler.
- (b) Simple physiological tests such as blood group tests.
- (c) Ethnological and anthropological studies.

P.3 To investigate environmental and developmental personality factors that may correlate significantly with P.K. ('S'2).

- (a) Age groups, according to chronological age and also according to appropriate tests.
- (b) Any tests measuring developmental and environmental personality factors.
- (c) Sociological studies.

P.4 To investigate any physiological changes which may be detectable when a subject attempts P.K. Broadly speaking correlations

The possibility is either based on or connected to the evidence or speculations indicated in brackets.

are possible between (i) physiological changes and 'on' and 'off' (or P.K. and non-P.K.) patterns, i.e. directional success is not necessary, (ii) between physiological changes and single P.K. scores. (P.K. strength and P.K. directions may provide the scores). ('S'4)

- (a) Electro-encephalogram. (45)
- (b) 'Lie detection' techniques. (46)
- (c) Any other tests that will measure any other physiological changes which may be accompanying P.K. output.

P.5 To standardise motivation ('C'3).

- (a) By playing actual games (using dice). (47)
- (b) By using stimuli which are similar to different subjects or which can be adjusted sufficiently to be similar.

For instances: electric shocks.

P.6 To standardise Experimenter-subject relationship. ('C'4)

- (a) By standardising instructions, etc.
- (b) By eliminating the experimenter in the actual test situation.

P.7 To pretend that a P.K. experiment is an orthodox or conventional physical or psychological test. ('C'5).

- (a) By explaining a P.K. test as a physical experiment, e.g.
 "When you wish something strongly, you are emotionally involved. If this happens your skin resistance changes. This change is measurable through two electrodes connected to your hand. The change in resistance produces a change in a current which flows through an electromagnetic

coil thus changing the strength of the magnetic field.

This coil will therefore attract the steel ball rolling from this lane into lane A instead of lane B."

- (b) By presenting strong stimuli, expecting that these stimuli will be strong enough to take care of the situation, e.g. everytime a ball rolls into lane A the subject receives an electric shock. It may be expected that the subject will not "wish" to receive a shock even if no further instructions are given.

P.8 To test differences between non-P.K. and P.K. rather than between P.K. (direction A) and P.K. (direction B). ('B'5, 'D'1, 'S'4).

P.9 To avoid changing of directions or changing of period (from non-P.K. to P.K.) too quickly, because a certain amount of 'carry over' may be expected. ('B'5), ('C'2).

- (a) By increasing, for instance, the number of dice thrown per run.
- (b) By increasing the number of unstable equilibrium positions per dice per run.
- (c) By using different objects to dice.
 - (i) microbalance. (48)
 - (ii) spinning coins. (49)
 - (iii) ball and right-left chance deviation.

P.10a and P10b may also be tried on P10c.

The latter case can be illustrated in the following way.

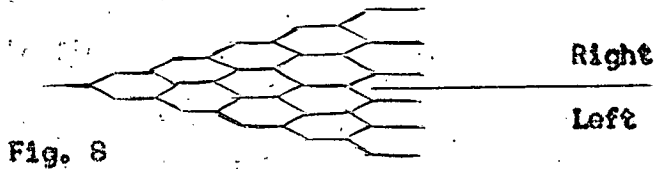


Fig. 8

P.11 To design experiments with analogue measurement rather than with digital measurement. Provided that the analogue measurement unit is small enough P.K. may be scored positively in cases where a digital system would show no positive results at all. ('S'5)

- (a) By measuring the time it takes an object to move between two fixed points.
- (b) By measuring spacial deviations of a moving object from a fixed line reference. (24)

P.12 To design experiments where the maximum physical forces that may be present near an unstable equilibrium position are measurable and comparatively small. (B'6)

For instance - (1) microbalance. (48)

(11) Millikan's Oil Drop Experiment. (50)

All the above possibilities may increase the reliability of secondary evidence ('B', 'C', 'D', evidence). If a secondary phenomena really exists but can be measured with a small reliability only, then the reliability is likely to improve when the number of subjects or the number of runs is increased.

For economic reasons, or reasons of general efficiency this is generally speaking, undesirable. However, if a new test is of

higher efficiency than previous ones, then the reliability may also be improved by working with more subjects than before, in this new test.

P.13 To improve the efficiency of the design and analysis of an experiment.

(a) By selecting successful subjects from the general public.

(1) through wireless instructions) motivation by competition
(11) Through television instructions) prizes etc.

(b) By selecting successful subjects, say from all first year psychology students in as many universities as possible, e.g. test student on E.S.P. and index results at some central university.

P.14 To improve the efficiency for an experimental design by complete automation (including the recording of data).

P.15 To make P.K. success conscious. ('B'3)

(a) By attempting to find subjects who are conscious of the scores.

(b) By stimulating the subject through drugs, electric shocks, etc.

There are some further possibilities which do not necessarily fall under the group of improving reliability and efficiency. If successful they would provide evidence of type 'B'4.

P.16 To increase the maximum distance over which P.K. can be scored.

P.17 To increase the maximum time in precognition experiments. (P.K. experiments may be tried too but are more difficult to design).

P.18 To vary size of P.K. objects in a systematic way and to attempt to find a maximum size.

P.19 To vary the type of P.K. objects systematically through different experiments.

(a) By varying surface structure.

(b) By varying the material (chemical or atomic structure).

(c) By varying from inanimate to animate 'objects'. (1)

P.20 To test whether non-human animals have P.K. or Psi abilities. (1)

P.21 To investigate normal (i.e. not experimental) situations where psi may have been active.

(a) Activities as homing of pigeons, pet dogs, etc.

(b) By investigating certain gambling scores which can be easily analysed.

(c) By investigating certain historical records. For example, the different results obtained in Millikan's Oil Drop experiment. Some attempts may be made to find whether these investigators had any strong opinions about the results before they started with the experiment and whether there exists any significant agreement between experimenter's expectations and actual result.

P.22 To investigate chance situations in general, e.g. is it possible to maintain 50/50 chance situation over a long period? What are the causes for change (in one direction) in this chance relationship? This may be called basic research.

P.23 To investigate the magnitudes of forces in a given chance situation. ('S'5) This may also be called basic research but may have direct applications. If it could be shown by physical

measurement that a die does not go through any unstable equilibrium positions with the angle between vertical line and centre of gravity $> \alpha$ (see page 31) during one run, then this run need not be scored.

- P.24 To investigate any phenomena that are of interest to psychical research groups.
- P.25 To investigate physical concomitants to psi activity which is not measurable in the subject but 'outside' him, e.g. change of temperature in the room in which the subject is working. (10)

Further Notes on these Possibilities.

P.5 Some attempts towards standardisation of motivations have been made. Rhine offered free tickets to the pictures and the playing of games was tried. (2,47) Also various drugs have been tried. (15) However, no system was tried yet which provided a direct reward or punishment for right or wrong P.K. scores. (51)

P.6 'Normal' standardisation in experimental situations has usually been a feature of most parapsychological experiments. However, experimenter-subject relationship seems to be a factor influencing psi results even when the 'normal' standardisation of instructions has been carried out. The only way to improve this further seems to be to eliminate the experimenters in the experimental situation altogether. This could be achieved by recording of instructions and complete automation of the experimental set-up (including the recording of data).

P.7 (see page 52)/

P.8 This type of investigation is most suitable for analogue measurements. ('P'11) For instance, if the time is measured it takes the

ball to roll between two fixed points, then if the experimental series is separated into P.K. and non-P.K. periods, one might expect the following graph ('D'1).

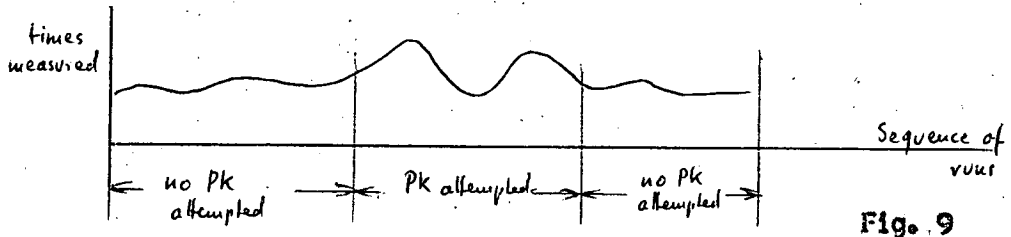


Fig. 9

But digital measurement (e.g. whether a ball goes to the right or to the left) may also be used if a reasonable number of actual runs are taken together. If, for instance, the first 30 runs without a subject (i.e. no P.K.) were 5/5, 5/5, 5/5, then with a subject (i.e. P.K. attempted) 2/8, 3/7, 7/3 and then again without a subject (i.e. no P.K.) 5/5, 5/5, 5/5, then subtracting right from left we would obtain the following nine scores which could be plotted and analysed in a similar way as the analogue measurements.

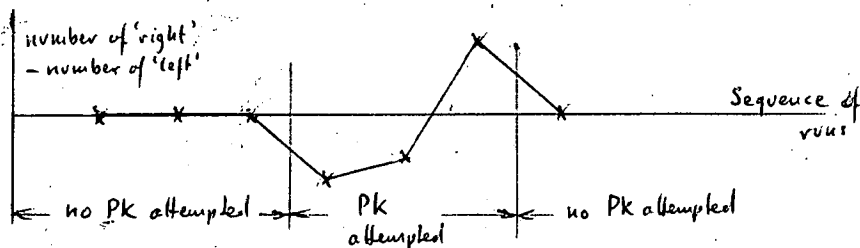


Fig. 10

P.9 This can best be illustrated by a graph similar to Fig. 9.

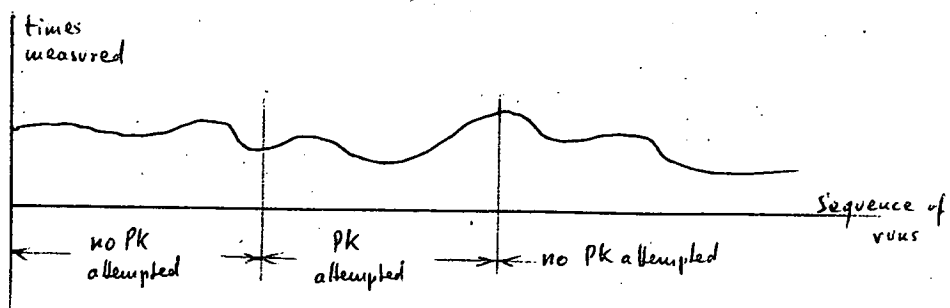


Fig. 11

It simply means that because of 'B'3, P.K. may start earlier and finish later than instructions may indicate.

Too many trials of the same type in one direction or simply P.K. (against non P.K.) will probably result in decreasing the right motivation. Therefore P.9 should be combined with P.5.

P.10 Mr. Forwald's experiment is an example of this possibility.

(24) Trying P10 in the left-right chance situations for balls (F.g 8) is technically difficult because it would take a considerable time to adjust the apparatus to obtain an approximate 50/50 deviation at each parting lane.

P.12 This was achieved with the microbalance and I will here only mention that the forces necessary to produce a complete shift (equivalent to turning a resting die over) were of known magnitude and comparatively small. (48)

P.14 This possibility seems very desirable since it will simplify and facilitate P.5, P.6, P.7, P.8, P.9, P.11 and others.

P.17 The maximum time for precognition may be compared with the maximum time during which posthypnotic suggestions will work.

P.19c It would here be desirable to attempt to define a hierarchy of animate objects. For instance one may start with virus continuing with bacteria and one may end up with higher animals or plants.

P.20 Systematic tests have been carried out with pigeons. (51)

But it may be possible to try rats and other animals in particular it may be possible to test such animals under very severe conditions. (42)

P.21b A certain betting system in Germany seems particularly suitable for evaluation. Each week people can choose 6 numbers from one to fifty and they will 'win' if they 'hit' the 'drawn' number combination. It would be interesting to see whether more people choose some of the correct numbers than could be expected theoretically. To find the theoretical distribution we would have to make some adjustments first. Some numbers are favoured more than others (e.g. 3, 7, 13 etc.) and the previous winning combination may also tend to upset the pure theoretical distribution. However, these changes could be estimated by analysing samples over several weeks. An analysis of this type would involve considerable work, and goodwill of the company would be needed to obtain old betting sheets. Suitable electronic counting machines seem necessary. However positive results would indicate that psi may be active in ordinary life when no conscious psi effort is made. It would support P.7.

P.21c This was suggested by Professor McAulay.

P.24 This has always been a reasonable possibility. The main difficulty is to guard against fraud. Since this is very difficult (there are many reports on cases where fraud was discovered or confessed under rather controlled conditions) it is probably not a good starting point for new investigators. It may be that the 'Zeitgeist' must be taken into account. If scientists refuse to even look at any such findings and if the co-operation of scientists is desirable then this may speak

also against such experiments. However it is possible that general parapsychological findings have brought about or are about to bring about some changes in attitude.

P.25 is less difficult than P.24. H.Price (10) found that during a sitting with a medium there was a considerable change of temperature in the room which he recorded automatically. Price was not satisfied that the experiment was sufficiently controlled to discard other explanations. However, such unintentional side effects may be, for the very reason that they are unintentional, more suitable for investigation.

Summary of Goals, Evidence and Possibilities

Goals:

1. To find Theory.
2. To provide more 'B'4 evidence.
3. To find lawful relationships.
4. To improve reliability.

Evidence:

A1	P.K. exists.	} psi exists (statistical)
A2	<u>E.S.P. exists.</u>	

B1 Results decline.

B2 Large differences in results.

B3 Psi unconscious.

B4 Positive findings.

B5 Replacement occurs.

B6 P.K. exists only when forces are suitable.

C1 P.K. space independent.

C2 P.K. time independent.

C3 P.K. depends on S's motivation.

C4 P.K. depends on E-S relation.

C5 P.K. depends on S's belief.

C6 P.K. depends on other personality factors.

D1 High tension of S produces strong P.K. but wrong directions.

Speculations:

S1 P.K. limited by hereditary factors (B2).

S2 P.K. limited by developmental and environmental factors.

- S3 P.K. universal if stimulated(B2), (C3), (C6).
- S4 P.K. may be sometimes non-directive (B2), (B5), (D1).
- S5 Unstable equilibrium positions are not present in all runs (B6).
- S6 Physiological concomitants exist in the subject (C6).

Possibilities:

- P1 Improve statistical design (A1).
- P2 Select by hereditary factors (S1).
- P3 Select by environmental and developmental factors (S2).
- P4 Find physiological concomitants. (S6).
- P5 Provide and standardise motivation (C3).
- P6 Standardise E-S relationship (C4).
- P7 Pretent P.K. is non-parapsychological (C5).
- P8 Investigate fluctuations, testing non-P.K. against P.K. periods (B5), (D1), (S4).
- P9 Do not change periods too quickly (B5), (C2).
- P10 Provide more than one chance situation per run (S5).
- P11 Try analogue measurements (S5).
- P12 Decrease forces involved (B6).
- P13 Increase efficiency of design and number of subjects.
- P14 Increase efficiency by complete automation.
- P15 Try to make P.K. conscious (B3).
- P16 Increase positive findings : maximum space independence.
- P17 Find maximum time independence.

- P18 Vary size of P.K. objects. Find maxima.
- P19 Vary type of P.K. objects. Include animate matter.
- P20 Investigate P.K. ability in non-human animals.
- P21 Investigate 'normal' situation where P.K. may have been active.
- P22 Basic research on chance situations.
- P23 Basic research on magnitude of forces.
- P24 Investigate particular cases (mediums).
- P25 Try to find unintentional physical P.K. concomitants 'outside of' S.

Possibility groups: No.1 to 15 mainly to improve reliability.

No.16 to 25 mainly to extend the field.

Further grouping is possible, e.g. subject-related possibilities and apparatus related possibilities.

Appendix 4

An apparatus for very short projection times

Reprinted from Australian Journal of Psychology, 1962,

Vol. 14, No. 1, pp. 39-40

AN APPARATUS FOR VERY SHORT PROJECTION
TIMES

H. H. J. KEIL
University of Tasmania

*Appendix 4 has
been removed for
copyright or
proprietary
reasons.*

Appendix No. 5

Greville correction applied to the multiple calls in the
disguised GESP group tests.

Trial No.	Target symbol No.					1^2	2^2	3^2	4^2	5^2	A	B	C
	1	2	3	4	5								
1	25	8	2	2	3	625	64	4	4	9	706	141.2	77.2
2	1	15	17	5	2	1	225	289	25	4	544	103.8	44.8
3	3	5	11	13	8	9	25	121	169	64	368	77.6	13.6
4	6	12	3	15	4	36	144	9	225	16	430	86.0	22.0
5	6	3	12	5	14	36	9	144	25	196	410	82.0	18.0
6	9	8	7	9	7	81	64	49	81	49	324	64.8	.8
7	7	10	8	12	3	49	100	64	144	9	366	73.2	9.2
8	3	9	14	9	5	9	81	196	81	25	392	78.4	14.4
9	4	14	5	15	2	16	196	25	225	4	466	93.2	29.2
10	7	12	7	11	3	49	144	49	121	9	372	74.4	10.4
11	8	9	12	3	8	64	81	144	9	64	362	72.4	8.4
12	7	9	9	10	5	49	81	81	100	25	336	67.2	3.2
13	6	6	13	7	8	36	36	169	49	64	354	70.8	6.8
14	5	9	4	15	7	25	81	16	225	49	396	79.2	15.2
15	8	10	14	4	4	64	100	196	16	16	392	78.4	14.4
16	5	13	5	10	7	25	169	25	100	49	368	73.6	9.6
17	10	14	5	6	5	100	196	25	36	25	382	76.4	12.4
18	9	8	9	8	6	81	64	81	64	36	326	65.2	1.2
19	11	7	9	8	5	121	49	81	64	25	340	68.0	4.0
20	8	7	8	12	5	64	49	64	144	25	346	69.2	5.2
21	11	11	8	6	4	121	121	64	36	16	358	71.6	7.6
22	6	13	10	7	4	36	169	100	49	16	370	74.0	10.0
23	5	6	8	15	6	25	36	64	225	36	386	77.2	13.2
24	10	4	11	10	5	100	16	121	100	25	362	72.4	8.4
25	6	9	10	5	10	36	81	100	25	100	342	68.4	4.4
26	7	6	13	10	4	49	36	169	100	16	370	74.0	10.0
27	5	9	9	9	8	25	81	81	81	64	332	66.4	2.4

28	8 9 2 10 11	64 81 4 100 121	370 74.0 10.0
29	8 4 6 14 8	64 16 36 196 64	376 75.2 11.2
30	4 6 15 9 6	16 36 225 81 36	394 78.8 14.8
31	11 4 13 9 3	121 16 169 81 9	396 79.2 15.2
32	7 18 5 8 2	49 324 25 64 4	466 93.2 29.2
33	9 8 10 8 5	81 64 100 64 25	334 66.8 2.8
34	7 10 4 8 11	49 100 16 64 121	350 70.0 3.0
35	4 11 6 6 13	16 121 36 36 169	378 75.6 11.6
36	6 9 9 5 11	36 81 81 25 121	344 68.8 4.8
37	8 10 8 8 6	64 100 64 64 36	328 65.6 1.6
38	6 14 7 8 5	36 196 49 64 25	370 74.0 9.0
39	9 11 6 10 4	81 121 36 100 16	354 70.8 6.8
40	7 13 6 5 9	49 169 36 25 81	360 72.0 8.0
41	12 11 5 10 2	144 121 25 100 4	394 78.8 14.8
42	9 7 11 5 8	81 49 121 25 64	340 68.0 4.0
43	7 7 14 8 4	49 49 196 64 16	374 74.8 10.8
44	6 7 8 12 7	36 49 64 144 49	342 68.4 4.4
45	4 13 10 9 4	16 169 100 81 16	382 76.4 12.4
46	13 13 5 7 2	169 169 25 49 4	416 83.2 19.2
47	4 12 16 7 1	16 144 256 49 1	466 93.2 29.2
48	5 6 11 7 11	25 36 121 49 121	352 70.4 6.4
49	5 8 6 12 8	25 64 36 144 64	333 66.6 2.6
50	8 7 11 8 6	64 49 121 64 36	334 66.8 2.8
51	9 7 12 6 6	81 49 144 36 36	346 69.2 5.2
52	4 8 8 16 4	16 64 64 256 16	416 83.2 19.2
53	4 6 8 13 9	16 36 64 169 81	366 73.2 9.2
54	6 12 5 9 8	36 144 25 81 64	350 70.0 6.0
55	7 5 15 3 10	49 25 225 9 100	408 81.6 17.6
56	6 7 11 9 7	36 49 121 81 49	336 67.2 3.2
57	7 9 12 7 5	49 81 144 49 25	348 69.6 5.6
58	3 11 10 13 3	9 121 100 169 9	408 81.6 17.6
59	8 11 14 7	64 121 196 49	430 86.0 22.0
60	11 10 4 14 1	121 100 16 196 1	434 86.8 22.8

61	14	7	6	6	7	196	49	36	36	49	366	73.2	9.2
62	6	11	13	5	5	36	121	169	25	25	376	75.2	11.2
63	6	8	9	8	9	36	64	81	64	81	326	65.2	1.2
64	6	7	4	14	9	36	49	16	196	81	378	75.6	11.6
65	5	10	9	8	8	25	100	81	64	64	334	66.8	2.8
66	10	6	5	10	9	100	36	25	100	81	342	68.4	4.4
67	3	5	15	11	6	9	25	225	121	36	416	83.2	19.2
68	8	10	9	9	4	64	100	81	81	16	342	68.4	4.4
69	4	10	8	14	4	16	100	64	196	16	392	78.4	14.4
70	6	15	8	5	6	36	225	64	25	36	386	77.2	13.2
71	11	7	9	8	5	121	49	81	64	25	340	68.0	4.0
72	5	7	11	10	7	25	49	121	100	49	344	68.8	4.8
73	4	6	13	13	4	16	36	169	169	16	406	81.2	17.2
74	7	7	5	13	8	49	49	25	169	64	356	71.2	7.2
75	4	6	13	8	9	16	36	169	64	81	366	73.2	9.2
76	12	11	9	8		144	121	81	64		410	82.0	18.0
77	4	12	16	5	3	16	144	256	25	9	450	90.0	26.0
78	5	7	14	7	7	25	49	196	49	49	368	73.6	9.6
79	6	9	7	15	3	36	81	49	225	9	400	80.0	16.0
80	3	14	14	2	7	9	196	196	4	49	454	90.8	26.8
81	8	17	5	4	6	64	289	25	16	36	430	86.0	22.0
82	4	10	11	11	4	16	100	121	121	16	374	74.8	10.8
83	6	5	11	15	3	36	25	121	225	9	416	83.2	19.2
84	6	8	9	9	8	36	64	81	81	64	326	65.2	1.2
85	4	12	9	5	9	16	144	81	25	81	347	69.4	5.4
86	10	8	8	11	3	100	64	64	121	9	358	71.6	7.6
87	4	12	9	9	6	16	144	81	81	36	358	71.6	7.6
88	6	10	8	9	7	36	100	64	81	49	330	66.0	2.0
89	2	14	11	9	4	4	196	121	81	16	418	83.6	19.6
90	8	7	12	3	10	64	49	144	9	100	366	73.2	9.2
91	10	13	6	8	3	100	169	36	64	9	378	75.6	11.6
92	7	11	9	9	4	49	121	81	81	16	348	69.6	5.6
93	5	6	12	9	8	25	36	144	81	64	350	70.0	6.0

94	6	8	5	11	10	36	64	25	121	100	346	69.2	9.2
95	5	7	8	9	11	25	49	64	81	121	340	68.0	4.0
96	7	11	6	7	9	49	121	36	49	81	336	67.2	3.2
97	9	9	10	10	2	81	81	100	100	4	366	73.2	9.2
98	8	8	9	9	6	64	64	81	81	36	326	65.2	1.2
99	9	9	4	9	9	81	81	16	81	81	340	68.0	4.0
100	11	5	9	6	9	121	25	81	36	81	344	68.8	4.8
Totals						6	9	9	8	6			
						9	1	0	7	1			
						4	5	4	5	0			
Grand Total						3998							
												1131.8	

Trial 1 to 100, total of column C = 1131.8

Testing for GESP (direct hits)

$$80/\sqrt{1131.8} = 2.38$$

$$CR = 2.38 ; p \leq 0.02$$

Trial 1 to 25, subtotal column C = 363.6

Trial 26 to 50, subtotal column C = 250.0

Trial 51 to 75, subtotal column C = 262.4

Trial 76 to 100, subtotal column C = 255.8

Testing the difference between the 1st and the 3rd quarter.

Trial 1 to 25 and trial 51 to 75 = 363.6 + 262.4 = 626.0

$$48/\sqrt{626.0} = 1.92$$

$$CR = 1.92 ; p < 0.05$$

Testing the difference between the combined 1st and 4th and the combined 2nd and 3rd quarter.

$$70 \sqrt{1131.8} = 2.08$$

$$CR = 2.08 ; p < 0.05$$

Appendix 5

Summary of individual results in undisguised GESP

group tests with 40 subjects

Subject Listed under these letters elsewhere (pp.220-221)	Direct Hits				Total	No. of calls per target No.		
	1-25	26-50	51-75	76-100		1,5,	2,3,4,	
A		8	7	4	7	26	20	80
	G	8	6	8	4	26	35	65
	E	5	7	6	8	26	31	69
	L	5	2	6	6	19	30	70
	W	10	4	4	7	25	39	61
B		5	9	5	6	25	27	73
	T	3	7	3	8	21	34	66
	K	3	5	4	9	21	47	53
C		1	8	5	7	21	30	70
D		8	2	5	5	20	31	69
	J	9	3	4	7	23	28	72
	M	11	5	4	3	23	28	72
E		3	4	5	9	21	30	69
	D	8	2	5	8	23	25	75
	R	6	3	6	5	20	32	68
F		5	6	7	4	22	29	71
G		6	6	1	4	17	37	63
H		11	5	4	3	23	22	78
I		5	7	6	2	20	38	62

J		3	4	5	6	18	55	48
	A	7	5	6	8	26	31	69
K		7	8	7	3	25	42	58
L		5	3	6	6	20	37	63
M		10	6	5	3	24	27	72
N		8	4	7	5	24	36	64
O		5	6	7	4	22	30	70
P		5	5	5	5	20	24	76
Q		3	2	4	6	15	21	79
R		5	6	4	6	21	27	73
S		5	8	4	8	25	35	65
T		7	7	2	6	22	27	73
	V	6	5	3	7	21	31	69
	H	8	4	3	4	19	35	68
U		4	8	5	6	23	38	62
	F	8	2	4	7	21	27	73
V		3	7	8	5	23	31	69
	C	5	3	8	7	23	39	61
W		7	3	6	4	20	45	55
	S	7	6	2	8	23	32	68
	Q	7	8	4	4	23	37	63
Total		245	208	197	230	880	1300	2698

Appendix 5

Disguised GESP group tests with 40 subjects.

Chi square analysis to test the difference between the number of subjects who made less or more end calls (and more or less middle calls) than expected and the number of subjects expected to make such calls.

End calls are calls for target symbols 1, 5, and middle calls are calls for target symbols 2, 3, 4.

	fo	fe	total
No. of subjects who made less end calls than expected	36	20	56
No. of subjects who made more end calls than expected	4	20	24
Total	40	40	80

$$\begin{aligned}
 \text{chi square} &= \frac{16^2}{20} + \frac{16^2}{20} \\
 &= 12.8 + 12.8 \\
 &= 25.6
 \end{aligned}$$

$$p < 0.001 ; df = 1$$

Appendix 6

Disguised and reinforced GESP experiment with 25 individual test sessions. Random number sequences used to determine the order of magazines and target slides.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
02	03	13	11	03	09	10	04	10	11	06	01	12	08	03
09	02	08	07	13	12	03	06	14	13	02	07	10	05	12
04	05	11	06	07	04	06	01	01	09	04	11	02	02	01
08	12	03	14	10	02	11	08	11	03	11	14	14	04	06
13	07	05	04	12	08	07	14	06	07	09	09	08	13	14
06	14	09	10	06	06	01	12	03	01	14	06	04	11	10
11	10	01	01	01	13	13	09	08	06	08	04	06	10	08

16	17	18	19	20	21	22	23	24	25
10	05	04	07	11	13	13	02	03	02
04	11	14	14	10	10	07	07	06	11
13	01	08	03	07	03	12	04	01	14
05	13	06	06	13	02	04	10	07	03
02	03	11	09	03	11	02	06	10	05
07	07	10	01	06	05	05	12	13	07
12	10	02	11	01	07	09	13	11	09

25 Individual test sessions of the disguised and reinforced GESP experiment

Four runs per test session

Subjects as listed:		direct hits					(-1) displacement					No. of calls per target								
pp. 339-40	220-1	1	2	3	4	Total	1	2	3	4	Total	1	2	3	4	5	1+5	2+3+4	No shock	shock
A.		7	9	4	6	26	4	6	2	9	21	1	38	32	25	4	5	95	13/43 (741)	> 10/57 (430)
S.		4	4	6	3	17	9	8	8	4	29	11	34	28	18	9	20	80	8/40 (480)	> 6/60 (240)
	G	7	4	4	2	17	3	13	7	7	30	13	26	25	19	17	30	70	4/44 (224)	< 6/56 (264)
F		6	5	4	1	16	8	5	3	8	24	11	26	25	23	15	26	74	2/36 (128)	> 2/64 (72)
	R	4	7	5	4	20	9	5	5	4	23	15	29	21	21	14	29	71	17/35 (455)	< 16/65 (560)
	(Y)	7	7	3	5	22	7	3	11	1	22	11	24	32	26	7	18	82	1/43 (57)	< 3/57 (129)
G		4	3	6	1	14	4	6	6	10	26	16	20	26	22	16	32	68	3/35 (195)	< 13/65 (455)
	C	4	3	5	6	18	7	3	9	3	22	18	22	20	31	9	27	73	2/35 (130)	< 9/65 (315)
	F	6	6	3	9	24	5	4	6	6	21	13	33	30	12	12	25	75	4/39 (244)	> 6/61 (234)
	H	5	4	7	9	25	7	1	5	3	16	20	20	31	17	12	32	68	3/37 (189)	< 7/63 (259)
S		8	4	8	6	26	5	5	4	6	20	18	34	19	20	9	27	73	3/44 (168)	< 4/56 (176)
Q		5	3	9	4	21	10	6	3	6	25	10	26	29	25	10	20	80	4/42 (232)	< 8/58 (336)
K		3	10	6	6	25	5	3	4	8	20	18	14	28	24	16	34	66	0/39	< 5/61 (195)
H		4	2	4	6	16	7	3	6	3	19	36	21	10	18	15	51	49	6/35 (390)	< 15/65 (525)
J		9	4	8	5	26	3	6	9	6	24	37	17	14	18	14	51	49	12/42 (696)	> 12/58 (504)
	D	7	6	8	5	26	8	6	4	6	24	6	28	23	24	19	25	75	5/44 (280)	> 3/56 (132)

Subjects (Y) and (Z) did not participate in other tests.

E	J	7	6	6	4	23	4	7	4	3	18	28	26	20	15	11	39	61	5/35 (325) > 4/65 (140)
	K	7	4	6	4	21	6	7	6	5	24	10	25	23	27	15	25	75	2/38 (124) > 3/62 (114)
	T	3	7	7	4	21	2	4	5	7	18	14	26	25	19	16	30	70	-
	W	6	6	9	7	28	7	9	4	7	27	21	18	25	12	24	45	55	5/50 (250) < 13/50 (650)
	E	9	6	4	6	25	5	6	2	5	18	13	30	19	23	15	28	72	7/36 (448) < 13/64 (468)
	(Z)	4	6	7	5	22	6	4	6	4	20	7	35	17	21	20	27	73	8/38 (496) < 12/62 (836)
	Q	3	3	8	7	21	4	7	7	5	23	17	17	22	28	16	33	67	3/40 (180) < 6/60 (240)
		9	5	1	8	23	4	6	5	6	21	13	24	21	23	19	32	68	2/39 (122) < 4/61 (156)
W		5	5	7	6	23	8	4	4	8	24	24	21	22	20	13	37	63	3/43 (171) < 5/57 (215)
Total		1	1	1	1		1	1	1	1		4	6	5	5	3			
		4	2	4	2		4	3	3	4		0	3	8	3	4			
		3	9	5	9	546	7	7	5	0	559	1	4	7	1	7	748	1752	

Subjects (Y) and (Z) did not participate in other tests.

Appendix 6

Disguised and reinforced individual GESP tests with 24 subjects.

Chi square analysis to test the difference between the number of subjects who made less or more end calls (and more or less middle calls) than expected and the number of subjects expected to make such calls.

End calls are calls for target symbols 1, 5 and middle calls are calls for target symbols 2, 3, 4.

	fo	fe	total
No. of subjects who made less end calls than expected	21	12	33
No. of subjects who made more end calls than expected	3	12	15
Total	24	24	48

$$\begin{aligned}
 \text{chi square} &= \frac{9^2}{12} + \frac{9^2}{12} \\
 &= 6.75 + 6.75 \\
 &= 13.5
 \end{aligned}$$

$$p < 0.001 ; df = 1$$

Appendix 6

Disguised and reinforced individual GESP tests with 24 subjects.
 Chi square analysis to test the difference between the number
 of subjects who repeated target symbols previously shocked
 more often than previously non-shocked symbols (and the number
 of subjects with reversed repetition tendencies), and the
 number of subjects expected to make such repetitions.

Two test sessions in which one subject participated.

	No shock	shock
Session 1	8/40	6/60
Session 2	3/44	4/56

$$8/40 + 3/44 > 6/60 + 4/56$$

	fo	fe	total
No. of subjects who repeated more often after shock	15	11.5	26.5
No. of subjects who repeated more often after non-shock	8	11.5	19.5
total	23	23	46

$$\begin{aligned}
 \text{chi square} &= \frac{3.5^2}{11.5} + \frac{3.5^2}{11.5} \\
 &= 1.06 + 1.06 \\
 &= 2.12
 \end{aligned}$$

$$p < 0.2 \quad ; \quad df = 1$$

Appendix 6

**Some adaptations of standard equipment for automatic
psychological recordings**

**Reprinted from Australian Journal of Psychology, 1961,
Vol.13, No. 1, pp. 99-103.**

***Appendix 6
has been removed
for copyright or
proprietary reasons.***

Random Analysis

Gap Analysis Binominal Analysis

Same micrometer setting, 1000 automatic trials,

15.3.'61, 16.3.'61. Temperature controlled.

Numbers in brackets refer to the number of 'right and 'left' paths per 10 trials. These are used for the test for randomness based on the binominal distribution.

The other numbers tabulated indicate the amount of gap between left and right paths respectively.

Total No. of right paths 558.

Total No. of left paths 442.

Length of 'right' trials Length of 'left' trials

	100	1	
(6)	1 5	1 1 2	(4)
(4)	1 1 1 1	1 2 1 2	(6)
(5)	2 1 2	1 4	(5)
3(1 3 1	1 3 1	(5)
(6)	1 3 2	2 1 1	(4)
3(1 3	5 1	(6)

Length of 'right' trials Length of 'left' trials

	100	1	
(7)	1 6	2 1	(3)
(7)	5 2	1 2	(3)
(5)	2 1 1 1	1 2 2	(5)
7(6 1	3	(3)
(4)	1 1 1 1	1 2 1 2	(6)
(6)	2 2 2	1 2	(4)
4(2 2 2	1 1 2	(4)

Length of 'right' trials Length of 'left' trials

100	1	
<u>3</u>	<u>1</u>)3
(5) 2	3	(5)
<u>1</u>	<u>1</u>	
(5) 1	1	(5)
2	3	
<u>1</u>	<u>2</u>	
(7) 6	2	(3)
2(1	1	
1	3	
(4) 1	1	(6)
2	2)5
(4) 1	3	(6)
3	3	
<u>1</u>	<u>3</u>	
(4) 3	1	(6)
<u>1</u>	<u>2</u>)4
(2) 1	2	(8)
1	2	
<u>1</u>	<u>4</u>)5
(7) 5	1	(3)
<u>1</u>	<u>1</u>	
1	1	
(6) 1	1	(4)
2	1	
<u>2</u>	<u>1</u>	
(9) 5	1	(1)
5(4		
1	2	(3)
(7) 5	1)2
<u>1</u>	<u>1</u>	
(8) 7	1	(2)
<u>1</u>	<u>1</u>	
4(3	1	
(8) 4	1	(2)
<u>1</u>	<u></u>	

Length of 'right' trials Length of 'left' trials

100	1	
<u>1</u>	<u>2</u>	
(3) 1	4	(7)
<u>1</u>	<u>1</u>	
3(2	2	
(7) 4	2	(3)
<u>1</u>	<u>2</u>	
4(3	2	
(6) 1	2	(4)
<u>2</u>	<u>2</u>	
8(6	2	(2)
(8) 2	2	
<u>2</u>	<u>2</u>	
(5) 3	2	(5)
<u>1</u>	<u>3</u>	
(5) 1	2	(5)
<u>4</u>	<u>3</u>)5
(5) 1	2	(5)
<u>4</u>	<u>5</u>	
(4) 1	1	(6)
<u>3</u>	<u>1</u>	
9(6	1	(3)
(7) 1	2)4
<u>2</u>	<u>2</u>	
(6) 4	1	(4)
<u>1</u>	<u>1</u>	
<u>5</u>	<u>1</u>	
(8) 2	1	(2)
<u>1</u>	<u>1</u>	
3(2	1	
(7) 1	1	(3)
3	1	
<u>1</u>	<u>1</u>	
3(2	1	
(6) 1	1	(4)
<u>3</u>	<u>2</u>	

(cont.)

Length of
'right' trials

Length of
'left' trials

100

1

(6) 1 1
3 3 (4)

5(2
(8) 3 2 (2)
5

(7) 5 1 (3)
1 1

7(1 1
(8) 6 1 (2)
2 1

2 2
(6) 1 1 (4)
1 1

4(2 1
(8) 2 1 (2)
3 1

(6) 5 1 (4)
1 1

2 5
(4) 1 3 (6)
3 1

2 4
(5) 1 2 (5)
1 1

2 1
(3) 1 3 (7)
1 1

2 3
(5) 1 1 (5)
1 1

2 1
(3) 1 3 (7)
1 1

2 3
(5) 1 1 (5)
1 1

2 2
(4) 1 1 (6)
2 4

2 1
(7) 1 1 (3)
1 1

3 3
(7) 1 1 (3)
3 1

3 3
(7) 1 1 (3)
3 1

3 3
(7) 1 1 (3)
3 1

3 3
(7) 1 1 (3)
3 1

Length of
'right' trials

Length of
'left' trials

100

1

(4) 3 5 (6)
3(1 1

(8) 2 2 (2)
6 2

(7) 3 1 (3)
3 1

(3) 1 2 (7)
2 3

3 1
(5) 1 1 (5)
1 1

3 1
(6) 1 1 (4)
1 1

1 1
(6) 1 1 (4)
1 1

2 1
(5) 1 1 (5)
1 1

1 1
(5) 1 1 (5)
1 1

1 1
(7) 1 1 (3)
1 1

1 1
(5) 1 1 (5)
1 1

1 1
(3) 1 1 (7)
1 1

1 1
(3) 1 1 (7)
1 1

1 1
(2) 1 1 (8)
1 1

1 1
(5) 1 1 (5)
1 1

1 1
(5) 1 1 (5)
1 1

1 1
(1) 1 1 (9)
1 1

1 1
(6) 1 1 (4)
1 1

1 1
(6) 1 1 (4)
1 1

16.3.61

(Cont.)

Length of 'right' trials Length of 'left' trials

	<u>100</u>	<u>1</u>	
2(1	2	
(5)	1	2	(5)
	2	1	
3(<u>1</u>	<u>3</u>	
(5)	1	2	(5)
	2	2	
(2)	2	2	(8)
	2	6	
(7)	1	1	(3)
	4	1	
(4)	2	3	(6)
	2	1	
(3)	1	2	(7)
	2	6	
(4)	1	2	(6)
	2	3	
4(<u>1</u>	<u>1</u>	
(6)	3	1	(4)
	1	3	
3(<u>2</u>	<u>1</u>	
(4)	1	4	(6)
	2	1	
(7)	1	1	(3)
	6	2	
(7)	1	1	(3)
	5	2	
4(<u>1</u>	<u>1</u>	
(8)	3	1	(2)
	3	1	
5(<u>2</u>	<u>2</u>	
(6)	3	2	(4)
	3	2	
(5)	1	1	(5)
	2	1	
	1	1	
	1	2	

Length of 'right' trials Length of 'left' trials

	<u>100</u>	<u>1</u>	
(3)	2	2	(7)
	1	5	
(5)	4	3	(5)
	1	2	
(7)	4	3	(3)
	3	1	
6(<u>3</u>	<u>1</u>	
(7)	2	1	(3)
	2	1	
(6)	3	2	(4)
	3	2	
	4	1	
(7)	1	1	(3)
	2	1	
3(<u>2</u>	<u>1</u>	
(6)	1	1	(4)
	3	3	
8(<u>2</u>	<u>2</u>	
(8)	6	2	(2)
	2	1	
	2	1	
(7)	5	1	(3)
	1	1	
(7)	2	1	(3)
	5	2	
	2	2	
(4)	2	1	(6)
	3	2	
(6)	1	2	(4)
	5	2	
(4)	2	4	(6)
	2	1	
	1	1	
(6)	2	2	(3)
	4	1	(4)
	1	1	
	1	4	
(3)	1	1	(7)
	1	2	

(cont.)

Length of 'right' trials Length of 'left' trials

	100	1	
(5)	1	1	(5)
	4	1	
		3	
(7)	6	1	(3)
	1	2	
	2	1	
(6)	3	1	(4)
	1	1	
		1	
(7)	4	1	(3)
	3	2	

Calculation of zero length (=0) of 'left' trials

0 = Total length of 'right' trial lengths n = 2 minus number
of 'right' trial lengths n = 2 intervals.

$$0 = 461 - 137$$

$$0 = 324$$

Calculation of the expected frequencies (f_e) for the various gap lengths.

$$f_e = p^2 (1 - p) N \quad p = 0.442$$

$$1-p = 0.558$$

$$\begin{aligned} \text{e.g. } f_e \text{ for length 2} &= 0.442^2 \times 0.558 \times 558 \\ &= 60.83 \end{aligned}$$

(cont.)

n = length of 'left' trials.

	f_o	f_e	$f_o - f_e$	$(f_o - f_e)^2$	$(f_o - f_e)^2 / f_e$
0	324	311.36	12.64	159.77	.51
1	129	137.62	8.62	74.30	.54
2	56	60.83	4.83	23.33	.38
3	26	26.89	.89	.79	.03
4	9	11.88	2.88	8.29	.70
> 4	14	9.42	4.58	20.98	2.23
Total	558	558 = N			4.39 = χ^2

chi square = 4.39

for df = 5 p = 0.5

Binominal Analysis

Calculation of expected frequencies (fe)

r	1						
per 10 trials		a	b=0.558 ^r	c= 0.442 ¹	d	fe=axbxcxd	fe
0	10	100	1	0.0002846	1	0.02846	9.55
1	9	100	0.558	0.0006439	10	0.3595	
2	8	100	0.3114	0.001457	45	2.0417	
3	7	100	0.1737	0.003296	120	6.8702	
4	6	100	0.09695	0.007456	210	15.1800	15.2
5	5	100	0.05410	0.01687	252	22.9992	23.0
6	4	100	0.03019	0.03817	210	24.1994	24.2
7	3	100	0.01684	0.08635	120	17.4496	17.5
8	2	100	0.009398	0.195	45	8.2645	10.9
9	1	100	0.005245	0.442	10	2.31629	
10	0	100	0.002926	1	1	0.2926	
Total = 100.003		Total 35		100.003		100.35	

r 1							
distribution							
per 10 trials		fo	fo	fe	(fo-fe)	(fo-fe) ²	(fo-fe) ² /fe
0	10						
1	9	1	11	10.9	.1	.01	0.00
2	8	3					
3	7	7					
4	6	15	15	17.5	5.5	30.25	1.73
5	5	20	20	24.2	4.2	17.64	0.73
6	4	20	20	23.0	3.0	9.0	0.39
7	3	23	23	15.2	.2	.04	0.00
8	2	10	11	9.55	.45	.20	0.02
9	1	1					
10	0						

chi square = 2.87 at df = 5 $p \geq .7$

Appendix 8

Chi square analysis to test the difference between subject-trials and machine-trials in 25 test sessions.

	right		left		total
	fo	fe	fo	fe	
Total machine-trials	1597	1547	903	953	2500
Total subject-trials	730	780	530	480	1260
Total	2327		1433		3760

$$\text{chi square} = \frac{50^2}{1547} + \frac{50^2}{953} + \frac{50^2}{780} + \frac{50^2}{480}$$

$$= 1.62 + 2.62 + 3.21 + 5.21$$

$$= 12.66, \quad p < 0.001; \quad df = 1$$

Chi square analysis to test the difference between subject-trials and machine-trials in 26 test sessions, i.e. the test session with the extreme distribution is included.

	right		left		total
	fo	fe	fo	fe	
Total machine-trials	1682	1634	918	966	2600
Total subject-trials	775	823	535	487	1310
Total	2457		1453		3910

$$\text{chi square} = \frac{48^2}{1634} + \frac{48^2}{966} + \frac{48^2}{823} + \frac{48^2}{487}$$

$$= 1.41 + 2.39 + 2.80 + 4.73$$

$$= 11.33; \quad p < 0.001; \quad df = 1$$

Appendix 8

Chi square calculation to test the difference between the total 'fore' machine trials and 'post' machine trials.

	right		left		
	fo	fe	fo	fe	total
fore trials	800	798.5	450	451.5	1250
post trials	797	798.5	453	451.5	1250
Total	1597		903		2500

$$\text{chi square} = \frac{1.5^2}{798.5} + \frac{1.5^2}{798.5} + \frac{1.5^2}{451.5} + \frac{1.5^2}{451.5}$$

$$= 0.003 + 0.003 + 0.005 + 0.005$$

$$= 0.016$$

$$p \approx .9 \text{ for } df = 1$$

Appendix 8

Chi square calculations per test session for Tasmanian
PK-machine tests.

m - t = machine trials per test session

s - t = subject trials per test session

Subject		fo	fe	fo	fe	Total
A	m - t	54	49	46	51	100
	s - t	24	29	36	31	60
	Total	78		82		160

$$\text{chi square} = 0.51 + 0.49 + 0.66 + 0.81$$

$$= 2.67 ; p < 0.2$$

		fo	fe	fo	fe	Total
B	m - t	57	56	43	44	100
	s - t	26	27	24	23	50
	Total	83		67		150

$$\text{chi square} = 0.02 + 0.02 + 0.04 + 0.04$$

$$= 0.12 ;$$

		fo	fe	fo	fe	Total
C	m - t	64	64	36	36	100
	s - t	32	32	18	18	50
	Total	96		54		150

$$\text{chi square} = 0$$

Subject

	fo	fe	fo	fe	Total
m - t	58	53	42	47	100
s - t	22	27	28	23	50
Total	80		70		150

$$\begin{aligned}\text{chi square} &= 0.47 + 0.53 + 0.93 + 1.09 \\ &= 3.02 ; p < 0.1\end{aligned}$$

L

	fo	fe	fo	fe	Total
m - t	81	75	19	25	100
s - t	31	37	19	13	50
Total	112		38		150

$$\begin{aligned}\text{chi square} &= 0.48 + 1.44 + 0.97 + 2.77 \\ &= 5.66 ; p < 0.02\end{aligned}$$

N

	fo	fe	fo	fe	Total
m - t	81	75	19	24	100
s - t	33	38	17	12	50
Total	114		36		150

$$\begin{aligned}\text{ch square} &= 0.33 + 1.04 + 0.66 + 2.08 \\ &= 4.11 ; p < 0.05\end{aligned}$$

O

	fo	fe	fo	fe	Total
m - t	45	51	55	49	100
s - t	31	25	19	25	50
Total	76		74		150

$$\begin{aligned}\text{chi square} &= 0.71 + 0.73 + 1.44 + 1.44 \\ &= 4.32 ; p < 0.05\end{aligned}$$

O

Subject

F

	fo	fe	fo	fe	Total
m - t	55	56	45	44	100

s - t	29	28	21	22	50
-------	----	----	----	----	----

Total	84		66		150
-------	----	--	----	--	-----

$$\begin{aligned}\text{chi square} &= 0.02 + 0.02 + 0.04 + 0.03 \\ &= 0.13\end{aligned}$$

G

	fo	fe	fo	fe	Total
m - t	68	71	32	29	100

s - t	39	36	11	14	50
-------	----	----	----	----	----

Total	107		43		150
-------	-----	--	----	--	-----

$$\begin{aligned}\text{chi-square} &= 0.13 + 0.31 + 0.25 + 0.64 \\ &= 1.33\end{aligned}$$

H

	fo	fe	fo	fe	Total
m - t	62	61	38	39	100

s - t	29	30	21	20	50
-------	----	----	----	----	----

Total	91		59		150
-------	----	--	----	--	-----

$$\begin{aligned}\text{chi square} &= 0.02 + 0.03 + 0.03 + 0.03 \\ &= 0.13\end{aligned}$$

I

	fo	fe	fo	fe	Total
m - t	57	51	43	49	100

s - t	19	25	31	25	50
-------	----	----	----	----	----

Total	76		74		150
-------	----	--	----	--	-----

$$\begin{aligned}\text{chi square} &= 0.71 + 0.73 + 1.44 + 1.44 \\ &= 4.32 : p < 0.05\end{aligned}$$

Subject

J

	fo	fe	fo	fe	Total
m - t	56	49	44	51	100
s - t	18	25	32	25	50
Total	74		76		150

$$\text{chi square} = 1.00 + 0.96 + 1.96 + 1.96 \\ = 5.88 ; p < 0.02$$

V

	fo	fe	fo	fe	Total
m - t	41	41	59	59	100
s - t	20	20	30	30	50
Total	61		89		150

$$\text{chi square} = 0$$

W

	fo	fe	fo	fe	Total
m - t	72	74	28	26	100
s - t	39	37	11	13	50
Total	111		39		150

$$\text{chi square} = 0.05 + 0.15 + 0.11 + 0.31 \\ = 0.62$$

Q

	fo	fe	fo	fe	Total
m - t	58	59	42	41	100
s - t	31	30	19	20	50
Total	89		61		150

$$\text{chi square} = 0.02 + 0.02 + 0.03 + 0.05 \\ = 0.12$$

Subject

R

	fo	fe	fo	fe	Total
m - t	49	48	51	52	100
s - t	22	23	28	27	50
Total	71		79		150

$$\begin{aligned}\text{chi square} &= 0.02 + 0.02 + 0.04 + 0.04 \\ &= 0.12\end{aligned}$$

S

	fo	fe	fo	fe	Total
m - t	73	69	27	31	100
s - t	31	35	19	15	50
Total	104		46		150

$$\begin{aligned}\text{chi square} &= 0.23 + 0.52 + 0.46 + 1.07 \\ &= 2.28 ; p < 0.2\end{aligned}$$

T

	fo	fe	fo	fe	Total
m - t	72	70	28	30	100
s - t	33	35	17	15	50
Total	105		45		150

$$\begin{aligned}\text{chi square} &= 0.06 + 0.13 + 0.11 + 0.27 \\ &= 0.57\end{aligned}$$

U

	fo	fe	fo	fe	Total
m - t	72	66	28	34	100
s - t	27	33	23	17	50
Total	99		51		150

$$\begin{aligned}\text{chi square} &= 0.55 + 1.06 + 1.09 + 2.12 \\ &= 4.82 ; p < 0.05\end{aligned}$$

Subject

P

	fo	fe	fo	fe	Total
m - t	61	75	19	25	100
s - t	32	38	18	12	50
Total	113		37		150

$$\begin{aligned}\text{chi square} &= 0.46 + 1.44 + 0.95 + 3.00 \\ &= 5.87 ; p < 0.02\end{aligned}$$

P

	fo	fe	fo	fe	Total
m - t	65	70	35	30	100
s - t	40	35	10	15	50
Total	105		45		150

$$\begin{aligned}\text{chi square} &= 0.36 + 0.53 + 0.71 + 1.67 \\ &= 3.57 ; p < 0.1\end{aligned}$$

K

	fo	fe	fo	fe	Total
m - t	65	67	15	13	100
s - t	45	43	5	7	50
Total	110		20		150

$$\begin{aligned}\text{chi square} &= 0.05 + 0.38 + 0.09 + 0.57 \\ &= 1.09\end{aligned}$$

(Subject K)

This test session has been excluded from further analysis

as indicated on p. 218

Appendix 8

Chi square calculations to test the difference between each interval total of subject-trials and the complete total of machine trials.

Each interval total consists of the sum of 'right' and the sum of 'left' trials from corresponding intervals of 10 trials per 25 test sessions.

	right		left		
	fo	fe	fo	fe	total
Total machine-trials	1597	1589	903	911	2500
first interval total of subject-trials	151	159	99	91	250

$$\begin{aligned}
 \text{Total} & \quad 1748 \quad \quad 1002 \quad \quad 2750 \\
 \text{chi square} &= \frac{8^2}{1589} + \frac{8^2}{911} + \frac{8^2}{159} + \frac{8^2}{91} \\
 &= 0.04 + 0.40 + 0.07 + 0.70 \\
 &= 1.21 \quad \quad p < 0.3 \text{ for } df = 1
 \end{aligned}$$

	right		left		
	fo	fe	fo	fe	total
Total machine-trials	1597	1580	903	920	2500
second interval total of subject-trials	141	158	109	92	250

$$\begin{aligned}
 \text{Total} & \quad 1738 \quad \quad 1012 \quad \quad 2750 \\
 \text{chi square} &= \frac{17^2}{1580} + \frac{17^2}{920} + \frac{17^2}{158} + \frac{17^2}{92} \\
 &= 0.18 + 1.83 + 0.31 + 3.14 \\
 &= 5.46 \quad \quad p < 0.02 \text{ for } df = 1
 \end{aligned}$$

	right		left		total
	fo	fe	fo	fe	
total machine-trials	1597	1577	903	923	2500
third interval total of	138	158	112	92	250
subject-trials					
Total	1735		1015		2750

$$\text{chi square} = \frac{20^2}{1577} + \frac{20^2}{923} + \frac{20^2}{158} + \frac{20^2}{92}$$

$$= 0.25 + 0.43 + 4.35 + 2.53$$

$$= 7.56$$

$$p < 0.01 \text{ for } df = 1$$

Fourth interval total of subject-trials : the right-left distribution is the same as in the first interval.

Hence chi square = 1.21 ; $p < 0.3$ for $df = 1$

	right		left		total
	fo	fe	fo	fe	
Total machine-trials	1597	1584	903	916	2500
fifth interval total of	145	158	105	92	250
subject-trials					
Total	1742		1008		2750

$$\text{chi square} = \frac{13^2}{1584} + \frac{13^2}{916} + \frac{13^2}{158} + \frac{13^2}{92}$$

$$= 0.11 + 0.18 + 1.07 + 1.85$$

$$= 3.21$$

$$p < 0.1 \text{ for } df = 1$$

Appendix 8

Analysis of variance to compare the machine-trials with the subject-trials in the Tasmanian PK experiment.

Interval totals of left-trials only were obtained. Each interval total consists of the sum of left trials from corresponding intervals of 10 trials per 25 test sessions.

Test session trial No.	Interval totals of left-trials	
1-10	77)
11-20	94)
21-30	99)
31-40	91)
41-50	79)
51-60	99)
61-70	109)
71-80	112)
81-90	99)
91-100	105)
101-110	94)
111-120	83)
121-130	97)
131-140	86)
141-150	93)

'fore'

machine-trials : group 1

subject-trials : group 2

'post'

machine-trials : group 1

$$\text{group 1 mean} = \frac{903}{10} = 90.3$$

$$\text{group 2 mean} = \frac{524}{5} = 104.8$$

$$\text{gross mean} = \frac{903 + 524}{15} = 95.133$$

$$\begin{aligned} \text{Total sum of squares} &= 18.133^2 + 1.133^2 + 3.867^2 + \\ &4.133^2 + 6.133^2 + 1.133^2 + 12.133^2 + 1.867^2 + 9.133^2 + 2.133^2 \\ &3.867^2 + 13.867^2 + 16.867^2 + 3.867^2 + 9.867^2 \end{aligned}$$

$$\begin{aligned} \text{Total sum of squares} &= 328.805 + 1.284 + 14.954 + \\ &17.082 + 37.614 + 1.284 + 147.210 + 3.590 + 83.412 + \\ &4.550 + 14.953 + 192.294 + 284.500 + 14.954 + 97.358 \\ &= 1243.844 \end{aligned}$$

$$\begin{aligned} \text{Sum of squares within group 1} &= 13.3^2 + 3.7^2 + 8.7^2 + \\ &0.7^2 + 1.3^2 + 3.7^2 + 7.3^2 + 6.7^2 + 4.3^2 + 2.7^2 \\ &= 176.89 + 13.69 + 75.69 + 0.49 + 1.69 + 13.69 + \\ &53.29 + 44.89 + 18.49 + 7.29 \\ &= 406.10 \end{aligned}$$

$$\begin{aligned} \text{Sum of squares within group 2} &= 5.8^2 + 4.2^2 + 7.2^2 + \\ &5.8^2 + 0.2^2 \\ &= 33.64 + 17.64 + 51.84 + 33.64 + 0.04 \\ &= 136.8 \end{aligned}$$

$$\begin{aligned} \text{Total sum of squares within groups} &= 406.10 + 136.8 \\ &= 542.9 \end{aligned}$$

Sum of squares between groups

$$d_1 = 95.133 - 90.3$$

$$d_2 = 95.133 - 104.8$$

$$d_1^2 = 4.833^2$$

$$d_2^2 = 9.667^2$$

$$d_1^2 = 23.3589$$

$$d_2^2 = 93.5476$$

$$10d_1^2 = 233.579$$

$$5d_2^2 = 467.738$$

Total sum of squares between groups = 233.579 + 467.738

$$= 701.317$$

Control $542.9 + 701.317 = 1244.217 \approx 1243.844$

Source of variation	sum of squares	df	estimated variance
between groups	701.317	1	701.317
within groups	542.9	13	41.762

$$F = \frac{701.317}{41.762} = 16.793$$

$p = 0.01$ for $F = 9.07$ and $p = 0.001$ for $F = 17.81$

Appendix 9

Undisguised test with the PK apparatus at Duke

Subject		right target		left target	
		right	left	right	left
A	target right first	10	20	6	24
B	target left first	11	19	6	24
C	target right first	18	12	15	15
D	target left first	16	14	18	12
E	target right first	14	16	8	22
F	target right first	14	16	12	18
	total	83	97	65	115

Chi square analysis to test the difference between the right-left distribution when right target is wished for and the right-left distribution when the left target is wished for.

	right		left		total
	fo	fe	fo	fe	
target right	83	74	97	106	180
target left	65	74	115	106	180
total	148		212		360

$$\begin{aligned}
 \text{chi square} &= \frac{9^2}{74} + \frac{9^2}{74} + \frac{9^2}{106} + \frac{9^2}{106} \\
 &= 1.09 + 1.09 + 0.76 + 0.76 \\
 &= 3.70
 \end{aligned}$$

$$p < 0.55 ; df = 1$$

Appendix 10

PK dice tests with rotating cage.

Test sessions with Miss Shearer as subject are marked S

High aim target faces	Accumulated deviation at the end of successive intervals of 144 trials each. The end of the 5th interval coincides with the end of one test session.					Deviation within each interval.						
	Interval No.					Interval No.						
	1	2	3	4	5	1	2	3	4	5	Total	
Target 1	0	+2	0	-3	-4	0	+2	-2	-3	-1	-4	
	+3	-6	+3	-4	-7	+3	-9	+9	-7	-3	-7	
	+1	-4	-10	-5	0	+1	-5	-6	+5	+5	0	
	-7	-8	-4	-7	-1	-7	-1	+4	-3	+6	-1	
	S	-1	-5	-8	-8	-18	-1	-4	-3	0	-10	-18
Target 2	-2	0	+9	+5	+1	-2	+2	+9	-4	-4	+1	
	-2	+2	+5	+7	+8	-2	+4	+3	+2	+1	+8	
	-4	+3	+6	+15	+20	-4	+7	+3	+9	+3	+20	
	S	+2	+6	+3	+17	+18	+2	+4	-3	+14	+1	+18
	S	-3	-2	-6	-10	-10	-3	+1	-4	-4	0	-10
Target 3	+11	+5	+8	+12	+16	+11	-6	+3	+4	+4	+16	
	+2	+3	+7	+2	+6	+2	+1	+4	-5	+6	+8	
	-3	-3	-1	+8	+13	-3	0	+2	+9	+7	+13	
	S	+2	+4	+6	+5	+6	+2	+2	+2	-1	+1	+6
	S	-2	-4	+1	-6	-16	-2	-2	+5	-7	-10	-16
Target 4	-1	+5	+3	+8	+12	-1	+6	-2	+5	+4	+12	
	+5	+4	+8	+7	+14	+5	-1	+4	-1	+7	+14	
	+4	+6	+9	+14	+14	+4	+2	+3	+5	0	+14	
	-4	-4	+1	+4	+6	-4	0	+5	+3	+2	+6	
	S	-4	-5	-13	-13	-15	-4	-1	-8	0	-2	-15
Target 5	-6	-2	+2	+1	+2	-6	+4	+4	-1	+1	+2	
	+3	-3	0	+7	+10	+3	-6	+3	+7	+3	+10	
	+12	+9	+14	+15	+18	+12	-3	+5	+1	+3	+18	
	+3	+10	+9	+3	+6	+3	+7	-1	+6	+3	+6	
	S	-6	-13	-20	-26	-27	-6	-7	-7	-6	-1	-27

Target 6	+11	+6	+14	+22	+22	+11	-5	+8	+8	0	+22
	0	+11	+11	+16	+10	0	+11	0	+5	-6	+10
	+1	+1	+1	+4	+7	+1	0	0	+3	+3	+7
	+10	+9	+7	+5	+11	+10	-1	-2	-2	+6	+11
	-9	-5	+6	+6	+9	-9	+4	+11	0	+3	+9
Total						+16	+6	+49	+30	+34	135

Appendix 10

Low aim target face		Accumulated deviations at the end of successive intervals of 144 trials each.					Deviations within each interval											
No.		Signs as for high aim targets					Signs for high aim targets						Signs reversed					
		Interval No.					Interval No.					Total	Interval No.					Total
		1	2	3	4	5	1	2	3	4	5	Total	1	2	3	4	5	Total
1		-4	-12	-9	-4	-8	-4	-8	+3	+5	-4	-8	+4	+8	-3	-5	+4	+8
2	S	-4	-10	-4	-2	-7	-4	-6	+6	+2	-5	-7	+4	+6	-6	-2	+5	+7
3	S	+3	+8	+9	+10	+13	+3	+5	+1	+1	+3	+13	-3	-5	-1	-1	-3	-13
4	S	+3	+6	+8	+14	+9	+3	+3	+2	+6	-5	+9	-3	-3	-2	-6	+5	-9
5	S	+2	+1	+5	-1	-12	+2	-1	+4	-6	-11	-12	-2	+1	-4	+6	+11	+12
6		+3	+1	+5	+7	+8	+3	-2	+4	+2	+1	+8	-3	+2	-4	-2	-1	-8
Total													-3	+9	-20	-10	21	-3
High aim total													+16	+6	+49	+30	+34	+135
Grand total													+13	+15	+29	+20	+55	+132

Absolute total for Miss Shearer, i.e. $\sum |d|$ where d = deviation per test session, reached 151 for 11 test sessions, i.e. for 330 PK runs (1 PK run = 24 trials).

Appendix II

Copy

SHORT MEMORANDUM

(Written by W. Saleh)

Experiment by J. Keil and W. Saleh with Psychology Students
on May 15, 16, and 17, 1962.

It was late in the year. The seniors had their preoccupation with their exams, but the experimenters went ahead with the experiment because they could not have done so before. At the same time they realized this handicap.

Object

A study of physiological variables and ESP.

The influence on the manifestation of ESP when the subject is told that the deck of 25 cards will be in one pile, or in 5 piles of 5 cards each, put each one near the other.

This separation of the 5 piles was inspired by Dr. B. Onetto. E₂ had done an experiment with him, comparing the hits when the deck of 25 was in one pile, and when it was in 5 piles (DT₅) (separated by 5 rubber bands) but put one on top of the other. Dr. BO scored less on the DT₅. But later he came and said he would do better on DT₅ if the 5 piles were separated from each other - not put one over the other. He scored higher on the five piles separated this way.

So, if a sophisticated person like Dr. BO with two Ph. D.'s and 2 years of parapsychology, thinks so, the students would be more likely to think so too. But nothing was mentioned to them.

Number of Runs

We planned to do four runs with each subject. If one started with DT_1 with one subject, we started with DT_5 with the next subject.

And when we started with DT_5 in the first run, we followed it with DT_1 and then a DT_5 and so on, thus alternating them.

We asked for 20 subjects planning to make 80 runs, but only 15 signed up and 11 came.

Procedure

When the subject came, the experiment was explained to him. He was shown the first room where the Polygraph was, and how the deck of cards would be when in one and in five piles. Then he was taken to the second office where he sat on a comfortable reclining armchair (chaise longue).

The electrodes were attached to the subject for the S.R. and the Plethysmograph, and he was asked a first question with three possible answers.

"Do you think it is possible to guess more than chance?"

".. yes.

".. no.

" .. I don't know.

We did not use the word ESP, to avoid any "counter-reaction from this group of subjects. We wanted them to think for the answer.

The second question was, "do you prefer DT_1 or DT_5 ?" "Do you think you will score higher on DT_1 or DT_5 ?"

E_2 wrote the answers then left for the room where the Polygraph

was, to ... (would you please describe what you did?)

During that time E₁ tried to "ease" the subject from his possible tenseness. He turned off the big light and left on a small indirect one. He tried to make the subject as comfortable as possible and tried to soothe him by soft words, asking him to forget all these usual thoughts, to "relax" as much as possible and to let "it" come. And when "it" comes.

He had better listen to his unconscious, so to speak, or to what comes from there without reasoning or thinking of the number of times he called one symbol or the other. Just to roll them down as wildly as they come, without thinking at all of them.

Usually it did not take a long time because the explanation of the experiment and the experiment itself had to be done in one hour, for the next subject was due. (The Experimenters talked after the experiment, so that it might be more fruitful to spend more time on the relaxation, on a next experiment).

When the signal was given by Experimenter 2 that he was ready, Experimenter 1 said to the subject to start when he wants to and waited for him to start. He pressed a button at every call, that marked a dot on the polygraph record.

At the end of the four runs Experimenter 1 took the subject to the record room where Experimenter 2 was and checked the diagrams of the ST and Pleth. He said what he thought would be a hit and E2 marked them on a new record sheet.

After that, we checked the cells of the subject with the first record sheet. If the following subject was there, we postponed the results to another time. (This separation of the experimenters appeared later to be a handicap because when the subject moved part of his body (he was instructed not to, but sometimes could not help it), this could not be recorded on the diagram to enable the observer to separate this movement from the natural physiological reactions.

Results

As our statistician, Gaither, is in your part of the world (in Japan, India, etc.) I am giving you a copy of the total hits and misses in both GSR and Plethismograph.

Appendix 11

Physiological measurements during ESP tests.

A = Plethysmograph correctly predicted hits..

B = GSR correctly predicted hits..

C = ESP hits..

D = Total predicted hits per test session.

E = Calls per test session.

A/D	B/D	C/E
2/25	1/26	13/100
1/25	1/17	18/100
4/31	2/19	17/100
1/21	2/19	24/100
7/23	4/23	18/100
10/28	6/24	24/100
3/26	6/19	22/100
9/27	7/18	29/100
1/23	4/23	15/100
11/28	5/22	27/100
Total 49/257	38/210	207/1000

10 test sessions with physiological measurements.

Chi square analysis to test the difference between the number of correct and incorrect predictions (based on the plethysmographic recording) and the number of correct and incorrect calls.

	correct		incorrect		total
	fo	fe	fo	fe	
plethysmographic predictions	49	52	208	205	257
calls	207	204	793	796	1000
Total	256		1001		1257

$$\text{chi square} = 0.17 + 0.04 + 0.04 + 0.01$$

$$= 0.26$$

$$0.5 < p < 0.7$$

Chi square analysis to test the difference between the number of correct and incorrect predictions (based on the GSR recordings) and the number of correct and incorrect calls.

	correct		incorrect		total
	fo	fe	fo	fe	
GSR predictions	38	43	172	167	210
calls	207	202	793	796	1000
Total	245		965		1210

$$\text{chi square} = 0.58 + 0.12 + 10.15 + 0.03$$

$$= 0.88$$

$$0.3 < p < 0.5$$

For runs with 6 or more hits only.

A/D	B/D	C/E
1/5	1/7	6/25
1/6	1/3	6/25
2/10	1/5	6/25
1/5	1/5	12/25
2/6	2/5	8/25
10/28	6/24	24/100
2/12	3/9	15/50
6/18	5/12	24/75
0/5	1/6	7/25
7/14	4/11	19/50
Total	32/109	25/87
		127/425

Chi square analysis to test the difference between the number of correct and incorrect predictions (based on the plethysmographic recording) and the number of correct and incorrect calls (for runs with 6 or more hits only).

	correct		incorrect		
	fo	fe	fo	fe	total
plethysmographic predictions	32	32.5	77	76.5	109
calls	127	126.5	298	298.5	425
Total	159		375		534

chi square < 0.1

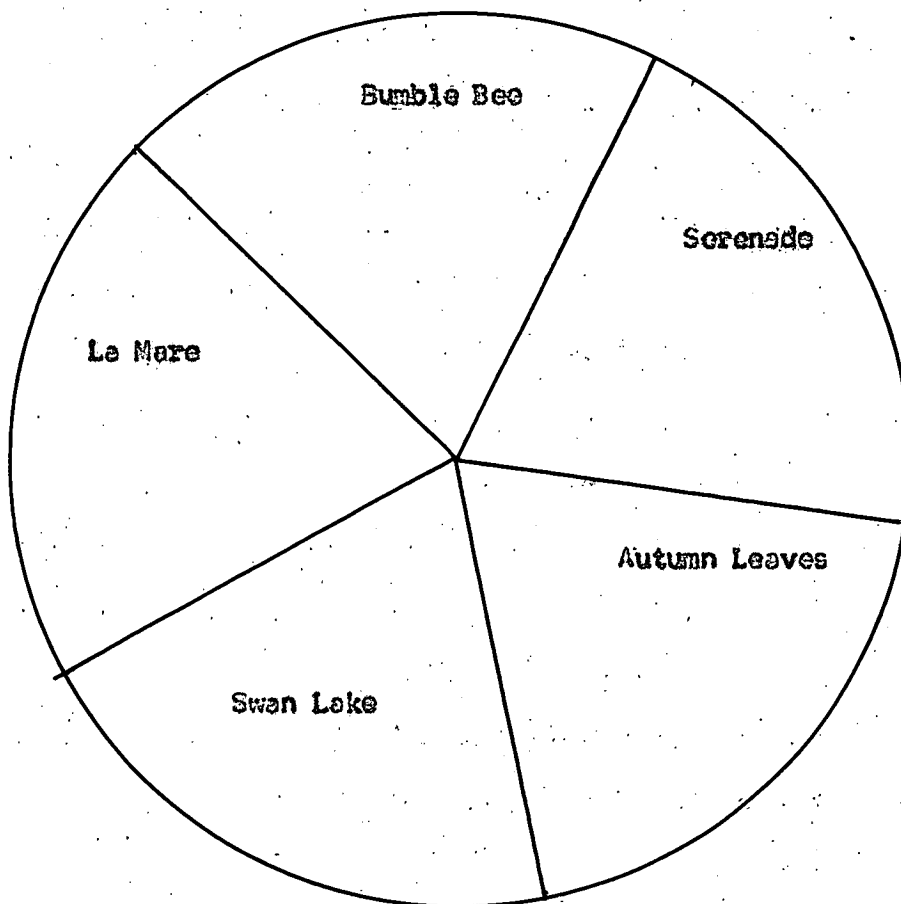
Chi square analysis to test the difference between the number of correct and incorrect predictions (based on the CSR recordings) and the number of correct and incorrect calls (for runs with 6 or more hits only).

	correct		incorrect		total
	fo	fe	fo	fe	
CSR predictions	23	25.8	62	61.2	87
calls	127	126.2	298	298.8	425
Total	152		360		512

chi square ≤ 0.1

Appendix 12

Example of target reminder sheet in GESP test with favourite music targets.



Appendix 12

GESP test with favourite music targets.

Total hits per interval of trials

	1-5	6-10	11-15	16-20	21-25	total deviat- ions	No. of runs
All test sessions	14	21	25	32	31	+23	20
'Low scoring' subjects	2	8	11	12	11	+3	3
'High scoring' subjects	12	13	14	20	20	+20	12
Single subject with highest scores	6	4	8	7	6	+11	4